NPSOR-93-007

NAVAL POSTGRADUATE SCHOOL

Monterey, California

STATISTICAL APPROACHES TO DETECTION AND QUANTIFICATION OF A TREND WITH RETURN-ON-INVESTMENT APPLICATION

Donald P. Gaver
Patricia A. Jacobs

December 1992

Prepared for:
NAVAIR (AIR 419)
Washington, D.C. 20550

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. W. West, Jr.
Superintendent

Harrison Shull
Provost

This report was prepared in conjunction with research funded by NAVAIR (AIR 419), Washington, D.C.

This report was prepared by:

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NPSOR-93-007	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL (If applicable) OR	7a. NAME OF MONITORING ORGANIZATION NAVAIR (AIR 419)
6c. ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code) Washington, D.C. 20550
8a. NAME OF FUNDING/SPONSORING ORGANIZATION NAVAIR (AIR 419) 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Washington, D.C. 20550	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Statistical Approaches to Detection and Quantification of a Trend with Return-on-Investment Application

12. PERSONAL AUTHOR(S)
Donald P. Gaver and Patricia A. Jacobs

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, month day)	15. PAGE COUNT
Technical	FROM______ TO______	1992 December	58

16. SUPPLEMENTARY NOTATION
The views expressed in this paper are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Changepoint problems, Maximum likelihood, Bayesian procedures, Cost of system upgrade

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Mathematical models are formulated for the possible onset and growth in subsystem degradation. The model recognizes that the time of onset of a degrading trend may be random, and hence initially unknown, and that the trend magnitude is also initially unknown. The trend magnitude will become better known as more data are accumulated. Maximum likelihood and Bayesian statistical procedures to estimate the time of onset and the trend magnitude are presented. A cost model is formulated to develop procedures (which recognize the uncertainty concerning the time of onset and trend magnitude) to determine estimated costs and the associated risks of upgrading the subsystem at different times in the future. Results of simulation studies of the procedures are presented.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT [x] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. [ ] DTIC USERS	21. ABSTRACT SECURITY CLASSICIATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL D. P. Gaver	22b. TELEPHONE (Include Area Code) (408) 646-2605 2c. OFFICE SYMBOL OR/Gv, OR/Jc

STATISTICAL APPROACHES TO DETECTION AND QUANTIFICATION OF A TREND WITH RETURN-ON-INVESTMENT APPLICATION

D. P. Gaver

P. A. Jacobs

1. INTRODUCTION

Members of a group of items such as aircraft, ships, tanks, etc use one or more subsystems of a particular type. These subsystems have basic design characteristics which result in specific values for measures of reliability, maintainability, repairability, etc. The values of these measures have a tendency to change, usually for the worse, over time. The time of onset of a degrading trend and the magnitude of the trend are unknown and must be estimated from data.

The evidence of the degradation of a subsystem suggests the possible economic and operational value of a subsystem upgrade, either by redesign of the existing subsystem or replacement with a new subsystem. The decision to upgrade a subsystem will, at least partially, be based on a comparison of the costs of remaining with the current subsystem and those of investing in the upgraded subsystem

The purpose of this report is to present preliminary models to assist in the assessment of the cost benefits of upgrading a subsystem in light of noisy data concerning its performance.

One part of the model is for the detection and qualitative description of a possible trend in noisy data. Two preliminary formal mathematical models are presented. One model is presented in Appendix A. Appendix A also describes maximum likelihood procedures to estimate the time of onset of system

degradation and the magnitude of the trend for one version of the general problem described. It turns out that this general problem type has been recognized early and studied by many under the name of **changepoint problems**; see Carlin, Gelfand and Smith (1992) for a very recent review of a certain style of approach, plus many references. The present exposition is self-contained and is directed specifically at an economic choice problem that potentially arises frequently in military logistics and procurement. A Bayesian model is presented in Appendix D. Appendix D also describes the Bayesian estimation procedure.

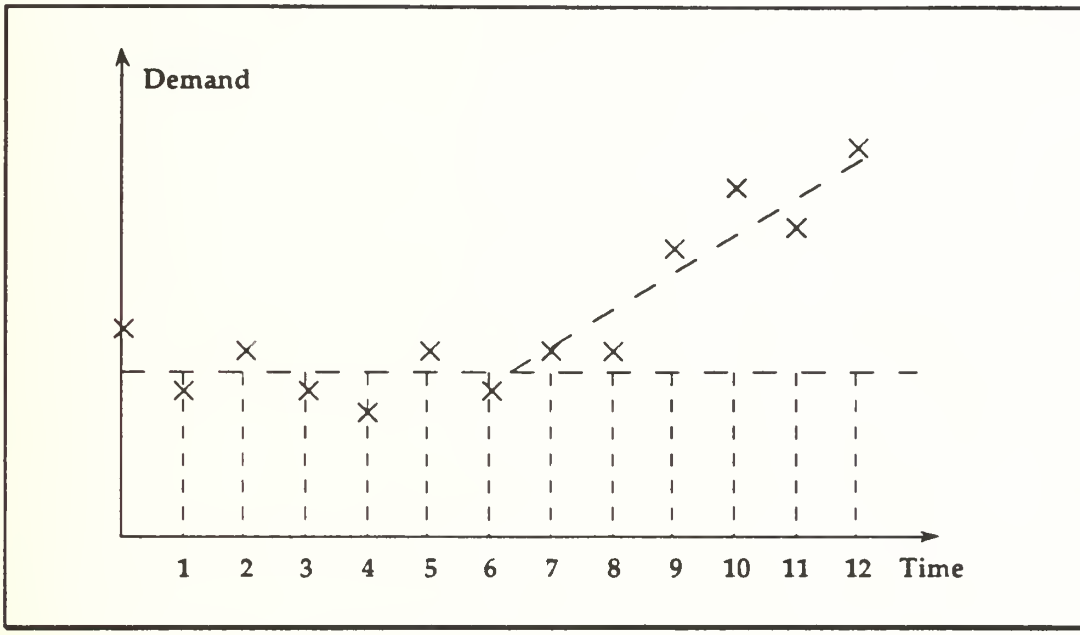
A second part of the overall model is a cost model which will have as input the estimates of the time of onset of system degradation and the magnitude of the trend. A simple cost model is presented in Appendix B. The cost model includes a fixed cost for upgrading the system as well as costs for each failure for both the current and upgraded subsystem.

Appendix C presents the results of a simulation experiment using simulated data to illustrate the type of information that can be obtained concerning the cost effectiveness of upgrading a subsystem using the model of Appendix A in the light of the uncertainty of the time of onset of system degradation and the magnitude of the degradation; the maximum likelihood procedure of Appendix A is used to estimate the parameters. Appendix D contains results of the Bayesian analysis of the same simulated data and compares the results of the two procedures.

In the following Section we informally discuss the general model in greater detail. In Section 3 we discuss an example and the results of the simulation experiments. Finally, in Section 4 we present conclusions.

2. THE MODEL

To give a concrete feel for the data that we are considering examine the following illustration. The ×'s represent actual data (e.g. total failures in a month), while the dotted lines represent the true, but hidden trend.



It is plain that there is little evidence of any change in the demand -level until possibly time t=9, when a retrospective look suggests that a change took place at about t=6 or 7. Successively more confirmation is given by observations at t=10 and 11, etc. The human eye picks up this new trend rather quickly (the human brain should remain skeptical about its permanence, and questioning concerning its magnitude and eventual level). _If_ the trend continues as suggested, greater and greater confirmation of its direction and magnitude becomes available; this is quantified by the preliminary mathematical models and statistical methods described in Appendices A and D. The statistical methods provide estimates of the true trend (denoted by the dotted line above) based on the number of failures observed (the ×'s above); the estimates are of the time of onset of system degradation and the magnitude of that degradation. As with the human eye, these

estimates become better as more confirmation is received concerning the permanence of the trend.

A unique feature of problem described in the Introduction is that the statistical models and estimation procedures are only part of the story. The problem also includes a comparison of the cost of remaining with the current subsystem and the cost of upgrading it. These costs will depend on the estimates of the time of onset of subsystem degradation and the magnitude of that degradation. Appendix B presents a simple cost model which depends on these estimates. The model is as follows. The current subsystem has a fixed cost for each failure. The upgraded system has a (large) known initial fixed cost for the upgrade and then a fixed cost for each failure. In this simple model it is assumed that the upgraded subsystem has a known fixed failure rate and a known cost of repair/replacement. Further, it is assumed that the subsystem can be upgraded in one time period. After the initial fixed cost, c_F, the mean cost of the upgraded system in each time period is a constant cost c_N multiplied by the (known) mean number of failures in each period. Informally, the cost model for the current subsystem is a (known) constant cost c_0 multiplying the true trend; (the dotted line in the above picture). However, since the true trend is unknown, for each time t an estimate of the cost of the current subsystem will be computed by multiplying c_0 by the estimate of the true trend obtained from data accumulated up to that time t. Estimates of the future cost of the current subsystem are computed by multiplying c_0 by the projected estimated trend. Thus, a decisionmaker will be comparing the (known) future cost of upgrading the subsystem with an estimated future cost of the current subsystem.

Since the estimates of the true trend have variability, the estimated future cost of the current system will also have variability. As with the estimates of the true trend, one can expect the estimates of future subsystem cost to be quite variable

until sometime after the onset of degradation; this variability is due to uncertainty in the estimates of the true trend and the time of onset. It is important to consider this uncertainty in the assessment of whether or not to upgrade the current subsystem. For example, it may be that the estimated mean future cost of the current subsystem is larger than that for the upgraded subsystem but that the uncertainty associated with the estimated mean future cost of the current subsystem is high. This may indicate that it is better to wait to accumulate more information concerning the apparent degrading trend before deciding to invest in the upgrade.

3. RESULTS

An example with simulated data is presented in Appendix C. The complete simulated data set appears in Figure 1. The true time of onset of degradation is at time 10. Before time 10 the true mean number of failures in each time period is $\mu=4$. After time 10 the true trend is linear with a slope of $\eta=1.5$. The true variability of the data about the true trend line is $\sigma^2=1$ for each time period.

The cost model for the example has the following features. The fixed cost per failure for the current subsystem is $c_0=2$. There is a fixed initial cost $c_F=225$ for upgrading the subsystem. The upgraded subsystem will have a lower mean number of failures, $\lambda=2$, in each time period but a higher cost per failure, $c_N=12$, than the current subsystem before the onset of degradation. After onset of degradation, the failure rate of the current subsystem may become larger than that of the upgraded subsystem due to the linear trend. The larger failure rate may offset the fixed cost of upgrading and make it economical to upgrade. There is also a time horizon, H=30, during which this subsystem or its upgrade will be used. If the onset of degradation is too close to the end of this time horizon, then because of the fixed upgrade cost, c_F, it will not be cost effective to upgrade the subsystem. The decision to upgrade depends on the estimates of the time of onset of system degradation and of the

magnitude of the trend. The assessment of the cost of upgrading should reflect the uncertainty of these estimates.

For each time $t \geq 5$ the following policies are considered: upgrade the subsystem at each future time until the time horizon; all potential upgrading times from the present time until the time horizon H are considered; that is, if the current time is t=11 then the policies that would upgrade the subsystem at time 11, time 12,..., time 29, (which is H-1), are considered. For each current time t, the (estimated) costs of these policies are compared to the (estimated) cost of never upgrading the subsystem. The "optimal" (minimum estimated mean cost) policy can then be found.

For comparison purposes the following is a description of the minimum mean cost policy in the (unrealistic) case in which the true trend in the data is known at each current time t. However, the decisionmaker is not omniscient and if the current time is before the onset of subsystem degradation, she will not know that this will occur. If at each time t, the correct trend for that time were known, then the minimum average cost policy would stay with the old subsystem until time 10. At time 10 (the time of onset of subsystem degradation), the decisionmaker instantly knows that change has occurred and the magnitude of the adverse trend so she can determine that the best policy is to upgrade the system at time 15. Suppose, however, that the decisionmaker becomes omniscient at a time after the time of onset of subsystem degradation (at time 10) and has not upgraded the subsystem. In this case for current times 11-15, the best policy is to upgrade the subsystem at time 15. For current times 16-23, the best policy is to upgrade the subsystem immediately. For current times greater than 23 the best policy is never to upgrade the subsystem (the cost of upgrading exceeds the advantage).

For each time $t \geq 5$, the simulation experiment of Appendix C considers the data accumulated up to time t and using the data as of that time estimates the time of onset of system degradation and the magnitude of the trend For each current time t, the estimated mean cost for each policy to upgrade the subsystem at some future time is computed using the current estimates of the trend.

Figure 2 presents the times to upgrade the subsystem which correspond to the minimum estimated mean cost policies for each current time. On the x-axis appears the "current" time. On the y-axis is the time to upgrade corresponding to the minimum estimated mean cost policy. If the minimum estimated mean cost policy is never to upgrade, then the time to upgrade is set equal to the horizon time, H=30. The following is a verbal description of these optimal policies. If the current time is either 5 or 6, the best policy is never to upgrade the system. Note that for the current time 7 the best policy based on the current estimates of the trend is to upgrade the system at time 9. An examination of the data in Figure 1 shows that around time 7 there is the local appearance of a positive slope. Hence, locally this policy is not unreasonable. However, the additional data point at time 8 results in updated estimates which indicate that the best policy at the current time 8 is never to upgrade. The best policy for current times 9-11 is never to upgrade. The best policy at current time 12 is to upgrade the subsystem at time 15. The best policy at time 13 is to upgrade at time 14. The best policy at time 14 is to upgrade at time 15. The best policy at time 15 is to upgrade immediately. The best policy at times 16-23 is to upgrade immediately. The best policy for current times larger than 23 is never to upgrade. Hence, except for current time 7, the optimal policy using the estimated projected future mean costs agree fairly well with the optimal policy for the case in which the true trend is known. This suggests that the estimation procedure of Appendix A requires some patience: one should not upgrade the system the first-

time that an upgrade is indicated, but let some time elapse for confirmation. One would not expect as close an agreement if the variability of the data about the true trend were larger since it would become more difficult to estimate the time of onset of the degradation and the magnitude of the degradation.

So far we have considered only point estimates of the mean cost of each policy. It may be that the variability of the estimated cost of each policy will yield more information. In Appendix C the variability of the mean policy costs computed using the estimation procedure of Appendix A is assessed using the computer intensive technique of bootstrapping. A brief description of the technique appears in that Appendix. Selected results are presented graphically in Figures 3-8. Each figure corresponds to a different current time and presents boxplots of bootstrap replications of the difference in estimated cost between a policy the upgrades at each future time and the policy that never upgrades. A description of the boxplot can be found in Appendix C; one can think of it as similiar to a very terse histogram. The y-axis represents the possible values of the cost differences. The x-axis represents the different possible times to upgrade. Since we are subtracting the cost of the policy of never upgrading from the cost of each policy which upgrades at a time in the future, a box and its appendages that correspond to negative values indicate that the estimated costs of the policy to never upgrade are higher than those for a policy that upgrades and hence it is better to upgrade the subsystem. Informally, the width of the box and its appendages are an indicator of the variability of the estimated cost differences between a policy which upgrades at a future time and the policy which never upgrades; the wider the box the more variable the estimated cost differences. One would expect the width of the boxes to be large around the time of onset of subsystem degradation because of the large uncertainty of the estimates. The circle in each box represents the mean; the line indicates the median.

Figure 3 presents boxplots of the mean cost differences for current time t=7 for all possible policies; the leftmost boxplot presents the cost differences for the policy of upgrading immediately at time 7; the next boxplot to the right presents the mean cost differences for the policy of upgrading the subsystem one time unit later, at time 8,etc. The boxplots of Figure 3 indicate that it is better to upgrade the subsystem almost immediately; the "best" time to upgrade is around time 9, but the sensitivity to the precise upgrade time is low. The spread (width of the boxes) of the cost differences is high and there appears to be not much difference between upgrading the subsystem immediately or waiting until time 14. Hence, the assessment of the variability of the estimated mean costs is providing the decisionmaker with the ability to compare different policies in light of the uncertainty of the estimates of the degrading trend. This additional comparison may prevent a premature decision to upgrade. Notice that bootstrapping does away with the instability that may result when the simple point estimate is used, i.e. Figure 2.

Suppose that the decision maker actually delays upgrading until later, either because she is still gathering information, she is concerned about the variability of the cost estimates, or because a rule tells her to wait. Figure 4 presents a similiar plot for the subsystem at current time 10; note that since the boxes of the cost differences correspond to positive values, the boxplots now indicate that the best policy is never to upgrade. Figure 5 presents boxplots for the policies evaluated at time 12; there is an indication that it is better to upgrade the subsystem; however, there appears to be little difference between upgrading at time 12 or at any time until time 17. Figure 6 presents boxplots for current time 15; there is a clear indication that one should upgrade the subsystem either immediately or in the next time period. Figure 7 presents the boxplots for the system at current time 18; there is an indication that one should upgrade immediately. Figure 8 presents the results for current time 25;

here the best policy is never to upgrade. Notice that as the decisionmaker accumulates more information concerning subsystem degradation, the estimates of the magnitude of the degradation are becoming better and the variability of the estimates for policy costs is becoming less.

Figures 9-11 present results of using the Bayesian analysis presented in Appendix D to obtain information concerning the cost effectiveness of switching to the new system. The same data set which illustrated the procedures in Appendix C is used; the data are presented in Figure 1. Since the bivariate normal has 5 parameters to be estimated, the estimation procedure begins with data $x_1,\ldots,x_6$.

Figure 9 presents the times to upgrade the subsystem which minimize the expected posterior mean cost for each decision time t=6,...29. On the x-axis appears the "current" time t. On the y-axis is the time to upgrade corresponding to the minimum expected mean cost policy using the posterior distribution given data $x_1,\ldots,x_t$. Comparison with Figure 2 indicates the following differences between the optimal Bayes policies and the optimal maximum likelihood (ml) policies presented in Appendix C. The Bayes policy for t=7 is still to upgrade; however, the time to upgrade is later (time 13) than the ml policy (which says to upgrade at time 9). The Bayes policy at time 12 is to upgrade at time 16 rather than time 15 for the ml policy. The Bayes policies and the ml policies are the same for the other times. Thus, the Bayes policy is more conservative than the maximum likelihood policy when there is a change in the policy from never upgrading to upgrading.

Simulation is used to obtain information concerning the variability of the posterior distribution of the average cost of each policy. For each current time t, a realization of the model is simulated from the posterior distribution and the average costs for each policy computed. Figures 10-11 present boxplots of 100 replications of the simulated difference in estimated cost between a policy that

upgrades at each future time and the policy that never upgrade. The y-axis represents the possible values of the cost differences. The x-axis represents the different possible times to switch.

Figure 10 presents the boxplots for simulated average policy cost differences using the posterior distribution at the current time of t=7. Figure 3 presents a similiar picture for bootstrap replications of estimated average policy cost differences using the maximum likelihood estimates of Appendix C at t=7. Comparison of the two figures indicates that the Bayes estimates of average cost difference are much larger and can be positive some of the time; recall that a positive difference implies that it is better never to upgrade. This behavior may be due to the Bayes procedure assessing greater variability to the estimated time of the onset of subsystem degradation than the procedure of Appendix A. Hence, the Bayes estimates are providing much less evidence of the need to upgrade. Recall that the onset of degradation does not occur until time 10. Hence, if the true model parameters were known, the best policy at time 7 (without omniscence) would be never to upgrade.

Figure 11 presents results for current time t=12. Comparing this figure with the corresponding maximum likelihood figure, Figure 5, indicates that there is little practical difference between the two procedures in this case also. However, the widths of the boxes for the Bayesian procedure are larger than those for the maximum likelihood procedure. The greater widths are an indication of greater uncertainty concerning the future average costs for each policy. As a result, the Bayesian procedure is providing less evidence of the need to upgrade. This greater variability is once again probably due to the Bayes procedure assessing greater uncertainty to the time of onset of degradation. This suggests that the Bayesian procedure may be more cautious than the maximum likelihood procedure.

4. CONCLUSIONS

Preliminary mathematical models have been formulated for the possible onset and growth in subsystem degradation. The model recognizes that the time of onset of a degrading trend may be random, and hence initially unknown, and that the trend magnitude is also initially unknown. The trend magnitude will become better known as more data is accumulated. Statistical procedures have been developed to estimate the time of onset and the trend magnitude. A rudimentary cost model has been used to develop procedures (which recognize the uncertainty concerning the time of onset and trend magnitude) to determine estimated costs and the associated risks of upgrading the subsystem at different times in the future. An experiment using simulated data gives reasonable results and indicates that the consideration of variability in policy costs due to uncertainty concerning the time of onset and trend magnitude can lead to wiser decisions.

REFERENCES

Carlin, B.P., A.E. Gelfand, and A.F.M. Smith, "Hierarchical Bayesian analysis of changepoint problems," **Applied Statistics** 41 (1992) pp.389-408.

Cobb, G.W. "The problem of the Nile: Conditional solution to a changepoint problem," **Biometrika** 65 (1978) pp. 243-252.

Efron, B. and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," **Statistical Science** 1 (1986) pp. 54-77.

Lewis, P.A.W. and L. Uribe, *The New Naval Postgraduate School Random Number Generator-LLRANDOMII,* Naval Postgraduate School Technical Report NPS 55-81-005, Monterey, CA.,1981.

Smith, A.F.M., "A Bayesian Approach to inference about a change-point in a sequence of random variables," **Biometrika** 62 (1975) pp. 407-416.

APPENDIX A

STATISTICAL MODELS AND ESTIMATION PROCEDURES

In this Appendix we present the statistical model and an estimation procedure.

1. MATHEMATICAL FORMULATION OF A STATISTICAL MODEL

We wish to use a sequence of fluctuating numerical values as a mathematical model for demand for a particular system during successive periods. Thus, consider a sequence of random variables with the following structure:

a) $X_1, X_2, X_3, \ldots X_C$ are identically and independently distributed, while

b) $X_{C+1}, X_{C+2}, \ldots X_t$ exhibit a linear trend. The time of onset of subsystem degeneration , C, called the **changepoint**, will realistically be unknown, as will the magnitude of the linear trend.

Example 1:

$$\begin{aligned} X_i &\sim N(\mu, \sigma^2), & 1 \le i \le C; \\ &\sim N(\mu + (i-C)\eta, \sigma^2), & C+1 \le i. \end{aligned} \tag{A.1}$$

This is shorthand for the assumption that X_i is normally/Gaussianly distributed with mean μ and variance σ^2 up to the changepoint time C, and is normally distributed thereafter, with variance σ^2 but with mean that grows—if the slope, η, is positive—linearly thereafter; μ is the mean number of failures in each time period before the onset of degradation; C is the time of onset of degradation; η is the slope of the linear trend after degradation; and the variance σ^2 is a measure of the variability of the actual number of failures about the true mean. This model should be appropriate for items whose mean demand/failure rate per time period, e.g. month, is reasonably large, but whose variance is relatively unchanged when and if a change in the mean occurs.

$$X_i \sim Po(\mu), \qquad 0 \le i \le C;$$

Example 2: (A.2)

$$\sim Po(\mu + (i - C)\eta), \quad C + 1 \le i$$

meaning that X_i is Poisson distributed with constant mean μ up to C, but thereafter has a linear trend. This model is most appropriate when the basic demand rate is small. Many other models are even more appropriate; for instance the Negative Binomial; Correlated or time-dependent demands may also occur. Attention to all of these is postponed. The most trustworthy model is likely to be based on some actual data. We are in the process of assembling such data.

2. LIKELIHOOD ESTIMATION

Suppose observations of the variables (numbers of failures during time periods 1,...,t) $X_1, X_2, \ldots, X_t$ are available; denote them by $x_1, x_2, \ldots x_t$. Then the likelihood function for the unknown parameters, μ, C, η, σ^2 is as follows for Example 1: since the number of failures in successive time periods are independent, for $1 \le C \le t$, and letting $data = (x_1, x_2, \ldots, x_t)$

$$L(\mu, C, \eta, \sigma^2; \text{data}) = \prod_{i=1}^{C} \frac{e^{-\frac{1}{2}(x_i - \mu)^2/\sigma^2}}{\sqrt{2\pi\sigma^2}} \prod_{i=C+1}^{t} \frac{e^{-\frac{1}{2}(x_i - \mu - (i-C)\eta)^2/\sigma^2}}{\sqrt{2\pi\sigma^2}} \tag{A.3}$$

so the log-likelihood is

$$l(\mu, C, \eta, \sigma^2; \text{data}) = \sum_{i=1}^{C} -\frac{1}{2}(x_i - \mu\ )^2 / \sigma^2 + \sum_{i=C+1}^{t} -\frac{1}{2}(x_i - \mu - (i - C)\eta)^2 / \sigma^2$$

$$-\frac{t}{2} \ln \sigma^2 + \text{constant.} \tag{A.4}$$

This can be concisely written as

$$l(\mu, C, \eta, \sigma^2; \text{data}) = \sum_{i=1}^{t} -\frac{1}{2}(x_i - \mu - (i - C)^+ \eta)^2 / \sigma^2 - \frac{t}{2} \ln \sigma^2 + \text{ constant.} \tag{A.5}$$

where

$$(i-C)^{+}=\begin{cases} i-C & \text{if } i \geq C, \text{ and} \\ 0 & \text{if } i < C. \end{cases} \tag{A.6}$$

Note that the above applies *if* there is a changepoint within the range of observation; otherwise, if $C > t$ then

$$L(\mu, C, \eta, \sigma^2; \text{ data}) = \prod_{i=1}^{t} \frac{e^{-\frac{1}{2}(x_i-\mu)^2/\sigma^2}}{\sqrt{2\pi\sigma^2}} \tag{A.7}$$

and

$$l(\mu, C, \eta, \sigma^2; \text{ data}) = \sum_{i=1}^{t} -\frac{1}{2}(x_i-\mu)^2/\sigma^2 - \frac{t}{2}\ln\sigma^2 + \text{ constant.} \tag{A.8}$$

Now in the following hold C fixed and behave as if it were known and the objective is to maximize l with respect to μ, η, and σ^2. Begin by differentiating with respect to μ:

$$\frac{\partial l}{\partial \mu} = \sum_{i=1}^{C} -(x_i-\mu)(-1)/\sigma^2 + \sum_{i=C+1}^{t} -(x_i-\mu-(i-C)\eta)(-1)/\sigma^2, \quad 0 \leq C \leq t; \tag{A.9}$$

$$= \sum_{i=1}^{t} -(x_i-\mu)(-1)/\sigma^2, \quad \text{if } C > t. \tag{A.10}$$

These expressions can be simplified and combined:

$$\frac{\partial \ell}{\partial \mu} = \frac{\left[t\bar{x}(t) - t\mu - \eta \sum_{j=1}^{t-C} j\right]}{\sigma^2}$$

$$= t\bar{x}(t) - t\mu - \left[\frac{(t-C)^2 + (t-C)}{2}\right]\eta \quad \text{for } t \geq C$$

$$= t\bar{x}(t) - t\mu \quad \text{for } t \leq C \qquad \text{(A.11)}$$

where $\bar{x}(t) = \sum_{i=1}^{t} x_i$

Rewrite this as

$$\frac{\partial l}{\partial \mu} = t[\bar{x}(t) - \mu] - \left[\frac{\left((t-C)^+\right)^2 + (t-C)^+}{2}\right]\eta \qquad \text{(A.12)}$$

where

$$(t-C)^+ = \begin{cases} t-C & \text{if } t \geq C; \\ 0 & \text{if } t \leq C. \end{cases} \qquad \text{(A.13)}$$

If the derivative is set equal to zero we obtain the first "normal equation"

$$\boxed{\mu + \psi_1(C,t)\eta = \bar{x}(t)} \qquad \text{(A.14)}$$

where here

$$\psi_1(C,t) = \frac{1}{t}\left[\frac{\left((t-C)^+\right)^2 + (t-C)^+}{2}\right] \qquad \text{(A.15)}$$

Next differentiate with respect to η: for t≥C:

$$\frac{\partial \ell}{\partial \eta} = \sum_{i=C+1}^{t} -(x_i - \mu - (i-C)\eta)(-(i-C))$$

$$= \sum_{i=C+1}^{t} x_i(i-C) - \mu \sum_{j=1}^{t-C} j - \eta \sum_{j=1}^{t-C} j^2 \tag{A.16}$$

$$= \sum_{i=C+1}^{t} x_i(i-C) - \mu t \psi_1(C,t) + \eta t \psi_2(C,t)$$

where

$$t\psi_2(C,t) = \sum_{j=1}^{t-C} j^2 = \frac{\left((t-C)^+\right)^3}{3} + \frac{\left((t-C)^+\right)^2}{2} + \frac{(t-C)^+}{6}. \tag{A.17}$$

Set the derivative equal to zero to obtain the second normal equation

$$\psi_1(C,t)\mu + \psi_2(C,t)\eta = \frac{1}{t}\sum_{i=C+1}^{t} x_i(i-C) \equiv \bar{x}_2(C,t) \tag{A.18}$$

Differentiate with respect to σ^2

$$\frac{\partial \ell}{\partial \sigma^2} = \sum_{i=1}^{t} -\frac{1}{2}\left(x_i - \mu - (i-C)^+\eta\right)^2\left(-\frac{1}{\left(\sigma^2\right)^2}\right) - \frac{t}{2}\frac{1}{\sigma^2}; \tag{A.19}$$

if this is equated to zero and solved for σ^2 there results

$$\hat{\sigma}^2 = \frac{1}{t}\sum_{i=1}^{t}\left(x_i - \hat{\mu} - (i-C)^+\hat{\eta}\right)^2. \tag{A.20}$$

Now solve the first two normal equations for the maximum likelihood estimate, conditional on C; the result is:

$$\hat{\mu}(C) = \frac{\psi_2\bar{x} - \psi_1\bar{x}_2}{\psi_2 - \psi_1^2} \tag{A.21}$$

$$\hat{\eta}(C) = \frac{\bar{x}_2 - \psi_1\bar{x}}{\psi_2 - \psi_1^2} \tag{A.22}$$

for $C<t$; for $C\geq t$, $\hat{\mu}(C)=\bar{x}$, $\hat{\eta}(C)=0$. These can now be substituted into (A.20) to obtain the maximum likelihood estimate for σ^2 in terms of the other estimates, all

conditional on the value of C. Finally substitute the above estimates into the expression for the negative of the log likelihood:

$$
\begin{aligned}
-\frac{2}{t}\ell\left(\hat{\mu}(C,t),C,\hat{\eta}(\mathrm{C,t}),\hat{\sigma}^2(C,t);\mathrm{data}\right) &\equiv S(C;\underline{\mathrm{data}}) \\
&= \frac{1}{t}\sum_{i=1}^{C}\frac{\left(x_i-\hat{\mu}(C,t)\right)^2}{\hat{\sigma}^2(C,t)}+\frac{1}{t}\sum_{i=C+1}^{t}\frac{\left(x_i-\hat{\mu}(C,t)-(i-C)\hat{\eta}(C,t)\right)^2}{\hat{\sigma}^2(C,t)} \\
&\quad + \ln\hat{\sigma}^2(C,t) \\
&= 1+\ln\hat{\sigma}^2(C,t)
\end{aligned}
\tag{A.23}
$$

and obtain the value of C that minimizes S(C;data) over the range (1,2,...,t); denote this by $\hat{C}(t)$; the last equality in the above expression follows from the definition of $\hat{\sigma}^2(C,t)$ given by (A.20). Thus, the estimate of C is chosen to minimize the sum of the squared residuals. If the minimum of S(C;data) occurs at t=C, then the conclusion is that no change has occurred in [0,t]. Note that all estimated parameter values, namely $\hat{\mu},\hat{\eta}$, and $\hat{\sigma}^2$ depend upon the C value in use, and so the dependence of S upon C involves that implicit dependency. Once $\hat{C}(t)$ is developed this value is substituted into the expression for $\hat{\mu},\hat{\eta}$, and $\hat{\sigma}^2$ to obtain the maximum likelihood estimates of those parameters. Note: there are other procedures for estimating the changepoint. One is explored in Appendix D. Others will be investigated in later work.

APPENDIX B

A MATHEMATICAL MODEL FOR COST AND RETURN-ON-INVESTMENT

The statistical model discussed in Appendix A is one part of the problem. In this Appendix we describe a simple cost model that links with the trend estimation procedure.

Suppose there is a cost c_0 incurred each time a subsystem fails; the total average cost incurred during the first t time periods is

$$C_0(t) = \begin{cases} c_0\mu(t+1) & \text{if } C \geq t, \\ c_0\sum_{s=0}^{t}[\mu + \eta(s-C)^+] & \text{if } C < t \\ = [c_0\mu + \eta\Psi_1(C,t)](t+1) & \end{cases} \tag{B.1}$$

where μ is the constant mean number of failure before the onset of subsystem degradation, C is the time at which the mean number of failures begins to show linear degradation, (trend), and η is the magnitude of that linear trend. Note that the parameters C, μ, and η are all unknown but may be estimated from data.

Suppose it is possible to upgrade the current system either by redesign of the existing system or replacement with a new system. The "new" system has a known constant mean number of failures in each time period λ and an average cost per unit failure of c_N. There is also a fixed cost c_F of changing to the new system. The total average cost of using the new system for t time units is

$$C_n(t) = c_F + c_N\lambda t \tag{B.2}$$

Assume there is a planning horizon H during which the parent system will be operative; when the horizon is reached all (remaining) parents are stored or disposed of. Note that we do not consider salvage costs in this treatment; they can be introduced if desired. At each time t one can compare the future cost of the

current system to that of the new system and choose to change to the new system if it has a lower mean future cost; that is, at time t the following decision can be made:

1) if $$C_0(H) - C_0(t) \le c_F + c_N \lambda (H - t) \tag{B.3}$$

then keep the current system;

2) if $$C_0(H) - C_0(t) > c_F + c_N \lambda (H - t) \tag{B.4}$$

then switch to the new system.

Since the parameters μ, C, and η are unknown, the estimation procedures described in the previous sections can be used to estimate them; the estimated average cost of the current system can then be obtained by computing $C_0(t)$ using the estimates. If desired, alternative estimation procedures can be utilized. There are many such, and selection should be based on tests with real data.

APPENDIX C

A SIMULATION EXPERIMENT

In this Appendix we describe a simulation experiment to illustrate uses of the model of Appendices A and B. The random numbers were generated using LLRANDOMII, cf, Lewis and Uribe (1981).

A data set of length 30 (e.g. 30 months) is generated from model (A.1) with parameters μ=4, η=1.5, σ^2=1, C=10; that is, data $x_1,\ldots,x_{30}$ are generated using (A.1) with the above parameters. The planning horizon is H=30. The cost per unit failure for the old system is c_0=2; the fixed cost for changing to the new system is c_F=225; the cost per unit failure for the new system is c_N=12; and the mean number of failures in each time period for the new system is λ=2. Thus the new system has a lower failure rate than does the old system initially, i.e. before degradation sets in at (unknown) time C, but the cost per failure for the new system is higher.

The following is a description of the optimal policy for each time t if the change to the new system has not occurred yet and the correct parameters are *known* for time t; if C has not yet occurred, then the policy decision uses only μ and not C or η for the decisionmaker is unaware that the system will degrade in the future. If at each time t, the correct parameters for that time were known, then the minimum average cost policy would stay with the old system until time 10. At time 10 (the changepoint), the decisionmaker instantly knows that change has occurred and the magnitude of the adverse trend so he can determine that the best policy is to change to the new system at time 15. Suppose, however, that the decisionmaker becomes omniscient at a time after the changepoint (at 10). For times 11-15, the best policy is to change to the new system at time 15. For times 16-23, the best policy is to change to the new system immediately. For times greater than 23 the best policy is never to change to the new system (the cost of change exceeds the advantage).

Now consider the decision maker in the simulated environment, with realistic deviation from the average trend. For each time t using the data $x_1,\ldots,x_t$ the following calculations are performed. Starting with t=5, estimates of μ,η,σ^2, and C using (A.20)-(A.22) and the procedure using (A.23) are obtained using noisy simulated data $x_1,\ldots,x_t$; denote the resulting estimates by $\hat{\mu}(t),\hat{\eta}(t),\hat{\sigma}^2(t)$, and $\hat{C}(t)$. The *estimated future mean cost* of a policy that switches to the new system τ time units in the future is computed for τ=0,...,30-t; that is,

$$\hat{C}_{new}(\tau;t)=\begin{cases} c_0\hat{\mu}(\tau+1)+c_F+c_N\lambda(30-(t+\tau)) & \text{if } \hat{C}>t \\ \left[\sum_{s=0}^{\tau} c_0(\hat{\mu}+\hat{\eta}(s+(t-\hat{C})^+))\right]+c_F+c_N\lambda(30-(t+\tau)) & \text{if } \hat{C}\le t \end{cases} \tag{C.1}$$

is computed for each τ and the minimum cost $\hat{C}_m(t)\equiv\min_{\tau\ge 0}\hat{C}_{new}(\tau;t)$ computed. This minimum cost is compared to the cost of doing nothing (and remaining with the current system) which is

$$\hat{C}_{old}(t)=\begin{cases} c_0\hat{\mu}(30-(t-1)) & \text{if } \hat{C}>t, \\ c_0(\hat{\mu}(30-(t-1))+\sum_{s=0}^{30-t}\hat{\eta}((t-C)^+ +s)) & \text{if } \hat{C}\le t \end{cases} \tag{C.2}$$

The policy associated with the minimum cost is chosen; in expressions (C.1) and (C.2) the estimates are $\hat{\mu}=\hat{\mu}(t)$, etc.,for time t.

Figure 1 presents the simulated data set. Figure 2 presents the optimal policies computed using (C.1) and (C.2) with the estimates using data $x_1,\ldots,x_t$ for each time t≥5; these computations assume that the decision maker is still getting information and that the change to the new system has not yet occurred. On the x-axis appears each time the best policy is computed. On the y-axis is the best policy's

time to change to the new system. The policy of never changing to the new system is represented by setting the time to change to the new system equal to the horizon H=30. The following is a description of the results. For times 5-6, the best policy is never to change to the new system. Note that at time 7 the best policy based on the current estimates is to change to the new system at time 9. An examination of the data in Figure 1 shows that around time 7 there is the local appearance of a positive slope. Hence, locally this policy is not unreasonable. However, the additional data point at time 8 results in updated estimates which indicate that the best policy at time 8 is never to change. The best policy for times 9-11 is never to change. The best policy at time 12 is to change to the new system at time 15. The best policy at time 13 is to change at time 14. The best policy at time 14 is to change at time 15. The best policy at time 15 is to change immediately. The best policy for times 16-23 is to change immediately. The best policy for times larger than 23 is never to change. This suggests that the current way of estimation requires some patience: one should not change to the new system the first-time that a change is indicated, but let some time elapse for confirmation.

Resampling or "Bayesian Bootstrapping"

A re-sampling technique called *the bootstrap* (cf. Efron et al. (1986)) can be used to assess the variability of the estimated mean cost associated with a policy due to the uncertainty of the parameter estimates. For each time t, 100 bootstrap replications $x_1(b;t),\ldots,x_t(b;t)$, b=1,...100, are generated using model (A.1) with the parameter estimates $\hat{\mu}(t), \hat{\eta}(t), \hat{\sigma}(t)$, and $\hat{C}(t)$. For each bootstrap replication the parameters μ, η, σ^2, and C are re-estimated obtaining $\hat{\mu}(b;t), \hat{\eta}(b;t), \hat{\sigma}(b;t)$, and $\hat{C}(b;t)$. For each bootstrap set of estimated parameters the future mean cost of a policy that switches to the new system τ time units into the future $\hat{C}_{new}(b;\tau;t)$ is computed using (B.1). The cost of never changing, $\hat{C}_{old}(b;t)$, is also computed using (B.2).

Figures 3-8 present boxplots for the differences in costs between the policy which says to change to the new system at time t+τ and the policy which says never to change, $\hat{C}_{new}(b;\tau;t) - \hat{C}_{old}(b;t)$, for τ=0,...,30-t for different times t for each bootstrap replication. The x-axis displays the time to change to the new system,t+τ, for each policy. The y-axis displays the cost differences. A negative value of the cost difference indicates that it is better to switch to the new system; the more negative, the greater the estimated mean advantage of changing to the new system.

All the graphical displays are produced by GRAFSTAT, a developmental product of IBM which the Naval Postgraduate School is using under a test agreement with IBM. The following description of the boxplot is taken from the documentation of GRAFSTAT. "The box portion of the plot extends from the lower quartile of the sample to the upper quartile. (The lower quartile is the point for which one quarter of the sample lies below and three quarters above. The upper quartile is analogous.) The line across the center of the box marks the median. The circle in the box represents the mean.

The distance from the lower to the upper quartile is called the interquartile distance and it will be represented by Q. The points at the ends of the two lines (called whiskers) are the smallest and largest points, respectively, within 1.5Q of the quantiles. The points beyond the whiskers are outlying values."

Figure 3 presents boxplots of the mean cost differences for t=7 for all possible policies; the leftmost boxplot presents the cost differences for the policy of changing immediately at time 7; the next boxplot to the right presents the mean cost differences for the policy of changing to the new system one time unit later, at time 8,etc. The boxplots of Figure 3 indicate that it is better to switch to the new system almost immediately; the "best" time is around t+τ=9, but the sensitivity to the precise change time is low. The spread of the cost differences is high and there

appears to be not much difference between switching to the new system immediately or waiting until time 14. Notice that bootstrapping does away with the instability that may result when the simple maximum likelihood estimate is used, i.e. Figure 2.

Suppose that the decision maker actually delays change until later, either because she is still gathering information or because a rule tells her to wait, and she agrees. Figure 4 presents a similiar plot for the system at time 10; note that since the cost differences are positive, the boxplots now indicate that the best policy is never to change. Figure 5 presents boxplots for the policies evaluated at time 12; there is an indication that it is better to switch to the new system; however, there appears to be little difference between switching at time 12 or at any time until time 17. Figure 6 presents boxplots for time 15; there is a clear indication that one should switch to the new system either immediately or in the next time period. The plot of Figure 7 presents the boxplots for the system at time 18; there is an indication that one should switch to the new system immediately. Figure 8 presents the results for time 25; here the best policy is to stay with the old system until the end of the time horizon H=30.

The boxplots can be interpreted as representing an approximate Bayesian posterior density for the true expected or mean cost, given observations up to time t. Their depth (length of box) becomes smaller as more data accumulates and uncertainty of estimation of the changepoint and the degradation rate, η, is reduced. But the depth of the boxes, plus the whiskers, provide perspective on the *risk* of changing soon, or waiting. Apparently the chance of making the wrong decision decreases if the decision maker waits, but also the value of making the more nearly correct decision decreases, for there is less time to the horizon.

It is important to be clear that the costs compared are estimated *mean* or *expected* costs, and not projected total costs, as there might be experienced during a future period. Boxplots that exhibit the probable range of these can also be exhibited. These more nearly represent true risk associated with actual return on investment.

Finally, note that all present calculations ultimately assume that the basic model is correct, or a good approximation. It may well be reasonable to check historical data for the approximate way in which degradation occurs -- it need not be a simple ramp of slope η, but perhaps precision of specification does not matter.

APPENDIX D

BAYES APPROACH TO CHANGEPOINT ECONOMICS

An enhanced version of the basic model presented in Appendix A is obtained by assuming that the changepoint (time of onset of degradation) is a random variable, C, with specified distribution whose parameter is unknown and subject to a probability density, $\pi(\cdot)$. Specifically, suppose

$$P\{C = k\} = (1-p)^{k-1} p, \tag{D.1}$$

i.e. is geometric, and that the prior $\pi(\cdot)$ is beta. As t advances one observes $x_1, x_2, \ldots, x_i, \ldots x_t$ and so in effect one has noisy observations on the outcomes of a biased coin flip with unknown success probability p.

We also generalize the normal model (A.1) to incorporate more general known trend functions than linear; in this Appendix, suppose that X-observations have the following structure:

a) $X_1, X_2, \ldots, X_C$ are independent and identically distributed normal random variables with mean μ and variance σ^2, while

b) $X_{C+1}, X_{C+2}, \ldots, X_t$ are independent with X_i normally distributed with variance σ^2 and mean $\mu + \eta g(i,C)$; that is,

$$\begin{aligned} X_i &\sim N(\mu, \sigma^2) && \text{if } 1 \le i \le C; \\ &\sim N(\mu + g(i,C)\eta, \sigma^2) && \text{if } C+1 \le i. \end{aligned} \tag{D.2}$$

The function g is an arbitrary nonnegative nondecreasing function with $g(i,C)=0$ for $i \le C$ representing the known form of the degrading trend which occurs after the time of onset of degradation. For the model of Appendices A and C, $g(i,C)=(i-C)^+$, a ramp starting at time C.

Putting (uninformative) priors on μ, η, and p, it is shown that the joint posterior density of these is straightforwardly obtained; the parameter σ^2 is initially

estimated from residuals without using a fully Bayes approach. In principle all of the above could be carried out for any arbitrary, but reasonable, discrete distribution that might better represent what is known about the changepoint process. A similiar statistical model was used by Smith(1975).

In what follows we sketch the development. Suppose that X-observations, $x_1,\ldots,x_t$, are available up to time t, it follows that

$$P\left\{p \in (dp), C=k, \mu \in (d\mu), \eta \in (d\eta), X_1 \in (dx_1),\ldots,X_j \in (dx_j),\ldots X_t \in (dx_t)\right\}$$

$$= \pi(p)(1-p)^{k-1} p \prod_{j=1}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x_j - \mu - \eta g(j,k))^2\right\} d\mu d\eta dp$$

for k≤t (D.3)

$$= \pi(p)(1-p)^{t} p \prod_{j=1}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(x_j - \mu)^2\right\} d\mu d\eta dp \qquad \text{for } k=t+1$$

The term involving $(1-p)^t$ represents the case in which no changepoint has occurred; we will set k=t+1 for this case.

By a completion of squares process one can write the likelihood function for given C=k as a bivariate normal density with parameters dependent on k and data up to t; the exponential term of the likelihood is written as

$$\prod_{j=1}^{t} \exp\left\{-\frac{1}{2\sigma^2}(x_j - \mu - \eta g(j,k))^2\right\}$$

$$= c \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(\mu-\bar{\mu})^2}{\gamma^2} - 2\rho\frac{(\mu-\bar{\mu})(\eta-\bar{\eta})}{\gamma\upsilon} + \frac{(\eta-\bar{\eta})^2}{\upsilon^2}\right) - K(k,t)\right\} \qquad \text{(D.4)}$$

for 1≤k≤t-1; for k≥t we have no changepoint so the exponential term of the likelihood is of the form

$$\prod_{j=1}^{t} \exp\left\{-\frac{1}{2\sigma^2}(x_j-\mu)^2\right\}$$
$$= c\ \exp\left\{-\frac{1}{2}\frac{(\mu-\overline{\mu})^2}{\gamma^2} - K(k,t)\right\} \tag{D.5}$$

where in the above c is a constant, and the parameters all depend upon k,t, and $\underline{x}(t)$, the data up to time t.

For k<t, the parameters of the bivariate normal (D.4) turn out to be

$$\overline{\mu}(k,t) = \frac{\overline{x}(t)\overline{g}_2 - \overline{x}_1(k,t)\overline{g}_1}{\overline{g}_2 - (\overline{g}_1)^2}; \tag{D.6}$$

$$\overline{\eta}(k,t) = \frac{\overline{x}_1(k,t) - \overline{g}_1\overline{\mu}(k,t)}{\overline{g}_2}; \tag{D.7}$$

$$\gamma^2(k,t) = \frac{\overline{g}_2}{\overline{g}_2 - (\overline{g}_1)^2}\frac{\sigma^2}{t}; \tag{D.8}$$

$$\upsilon^2(k,t) = \frac{1}{\overline{g}_2 - (\overline{g}_1)^2}\frac{\sigma^2}{t}; \tag{D.9}$$

and

$$\rho(k,t) = -\frac{\overline{g}_1}{\sqrt{\overline{g}_2}}; \tag{D.10}$$

where

$$\overline{x}(t) = \frac{1}{t}\sum_{j=1}^{t} x_j; \tag{D.11}$$

$$\overline{x}_1 \equiv \overline{x}_1(k,t) = \frac{1}{t}\sum_{j=1}^{t} x_j g(j,k) \tag{D.12}$$

$$\overline{g}_1 \equiv \overline{g}_1(k,t) = \frac{1}{t}\sum_{j=1}^{t} g(j,k); \tag{D.13}$$

$$\bar{g}_2 \equiv \bar{g}_2(k,t) = \frac{1}{t}\sum_{j=1}^{t} g^2(j,k); \tag{D.14}$$

and

$$K(k,t) = \frac{1}{2\sigma^2}\sum_{j=1}^{t}(x_j - \bar{\mu}(k,t) - \bar{\eta}(k,t)g(j,k))^2. \tag{D.15}$$

For the case k≥t

$$\bar{\mu}(k,t) = \frac{1}{t}\sum_{j=1}^{t} x_j; \tag{D.16}$$

$$\gamma^2(k,t) = \frac{\sigma^2}{t}; \tag{D.17}$$

and

$$K(k,t) = \frac{1}{2\sigma^2}\sum_{j=1}^{t}(x_j - \bar{\mu}(k,t))^2; \tag{D.18}$$

$\bar{\eta}(k,t) = 0$, $\upsilon^2(k,t) = 0$, and $\rho(k,t) = 0$. These values can be derived directly from (D.4) and (D.5); details are omitted.

If the bivariate normal form is utilized in (D.3) and the integration is the integration is performed over p we obtain the joint conditional density of C, μ, and η given the data and σ^2 in the form

$$P\{C = k, \mu \in (d\mu), \eta \in (d\eta) \mid \underline{x}(t), \sigma^2\}$$
$$= \pi^*(k,t)\frac{1}{2\pi(1-\rho^2)^{0.5}\gamma\upsilon}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(\mu-\bar{\mu})^2}{\gamma^2} - 2\rho\frac{(\mu-\bar{\mu})(\eta-\bar{\eta})}{\gamma\upsilon} + \frac{(\eta-\bar{\eta})^2}{\upsilon^2}\right]\right\} \tag{D.19}$$

for k<t where

$$\pi^*(k,t) = c^* 2\pi\sqrt{1-\rho^2}\,\gamma\upsilon\int_0^1 (1-p)^{k-1} p\pi(p)dp\exp(-K(k,t)); \tag{D.20}$$

for k≥t

$$P\{C = k, \mu \in (d\mu) \mid \underline{x}(t), \sigma^2\}$$
$$= \pi^*(k,t)\frac{1}{\sqrt{2\pi}\gamma}\exp\left\{\frac{1}{2\gamma^2}(\mu-\bar{\mu})^2\right\} \tag{D.21}$$

with

$$\pi^*(t,t) = c^*\sqrt{2\pi}\gamma\int_0^1 (1-p)^{t-1} p\pi(p)dp\exp(-K(t,t)) \tag{D.22}$$

$$\pi^*(t+1,t) = c^*\sqrt{2\pi}\gamma\int_0^1 (1-p)^t \pi(p)dp\exp(-K(t+1,t)) \tag{D.23}$$

and

$$c^* = [\sum_{k=1}^{t+1} \pi^*(k,t)]^{-1} \tag{D.24}$$

Note that {π*(k,t), k≤t} is the marginal probability that the changepoint occurs at any time k up to and including t; while π*(t+1,t) is the posterior probability that no changepoint has occurred up to time t; that is,

$$\pi(k,t) = P\{C = k \mid X_1 = x_1,\ldots,X_t = x_t\} \qquad \text{for } k\le t \tag{D.24a}$$

$$\pi(t+1,t) = P\{C > t \mid X_1 = x_1,\ldots,X_t = x_t\}. \tag{D.24b}$$

For each time t, the estimate of σ^2 is computed from the squared residuals for each possible value of C=k in the following manner; let

$$\hat{\sigma}^2(k,t) = \frac{1}{t-1}\sum_{j=1}^{t}(x_j - \overline{\mu}(k,t) - \overline{\eta}(k,t)g(j,k))^2 \qquad \text{if } k \le t \tag{D.25}$$

$$\hat{\sigma}^2(k,t) = \frac{1}{t-1}\sum_{j=1}^{t}(x_j - \overline{\mu}(k,t))^2 \qquad \text{if } k \ge t. \tag{D.26}$$

Finally, the estimate of the variance σ^2 based on data $x_1,\ldots,x_t$ is

$$\hat{\sigma}^2(t) = \sum_{k=1}^{t+1} \pi^*(k,t)\hat{\sigma}^2(k,t). \tag{D.27}$$

Given C=k, k<t, and the data $x_1,\ldots,x_t$, the posterior distribution of (μ,η) is bivariate normal with mean $(\overline{\mu}(k,t),\overline{\eta}(k,t))$, variance of μ equal to γ^2(k,t), variance of η equal to υ^2(k,t), and correlation ρ(k,t); for k=t,t+1, η=0 and the posterior distribution of μ is normal with mean $\overline{\mu}(k,t)$ and variance γ^2(k,t). Hence, given the data $x_1,\ldots,x_t$, the posterior distribution of (μ,η) is a mixture of bivariate normal distributions with mixture distribution {π*(k,t), k≤t+1}.

At time t, the future mean cost of a policy that switches to the new system τ time units in the future is given by (C.1) with $(t-C)^+$ replaced by g(t,C); the mean cost of remaining with the current system is given by (C.2) with $(t-C)^+$ replaced by g(t,C). Since the mean costs are linear in μ and η, the posterior distribution of the future mean cost given the data $x_1,\ldots,x_t$ is a mixture of normal distributions with mixture distribution $\{\pi^*(k;t),\ k\leq t+1\}$. The variability of the estimated future mean cost for each policy can be evaluated either by computing the percentiles of the posterior distribution or by simulating the posterior distribution.

Figures 9-11 present results of using the Bayesian analysis presented in this Section to obtain information concerning the cost effectiveness of switching to the new system. The same data set which illustrated the procedures based on maximum likelihood in Appendix C is used; the data are presented in Figure 1. Since the bivariate normal has 5 parameters to be estimated, the estimation procedure begins with data $x_1,\ldots,x_6$. The initial estimate of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{4}\sum_{j=1}^{5}(x_j - \bar{x})^2 \tag{D.28}$$

where $\bar{x}$ is the sample average of the first 5 data points. For each time t, estimates of the posterior distribution are obtained from equations (D.6)-(D.27). The updated estimate of σ^2 is used as input for the calculations for the next time period.

Figure 9 presents the times to upgrade the subsystem which minimize the expected posterior mean cost for each decision time t=6,...29. On the x-axis appears the "current" time t. On the y-axis is the time to upgrade corresponding to the minimum expected mean cost policy using the posterior distribution given data $x_1,\ldots,x_t$. Comparison with Figure 2 indicates the following differences between the optimal Bayes policies and the optimal maximum likelihood (ml) policies presented in Appendix C. The Bayes policy for t=7 is still to upgrade; however, the time to upgrade is later (time 13) than the ml policy (which says to upgrade at time 9). The

Bayes policy at time 12 is to upgrade at time 16 rather than time 15 for the ml policy. The Bayes policies and the ml policies are the same for the other times. Thus, the Bayes policy is more conservative than the maximum likelihood policy when there is a change in the policy from never upgrading to upgrading at some time.

Simulation is used to obtain information concerning the variability of the posterior distribution of the average cost of each policy. For each current time t, a realization of (C,μ,η) is simulated from the posterior distribution and the average costs for each policy computed. Figures 10-11 present boxplots of 100 replications of the simulated difference in estimated cost between a policy that switches to the new system at each future time and the policy that never switches. The y-axis represents the possible values of the cost differences. The x-axis represents the different possible times to switch.

Figure 10 presents the boxplots for simulated average policy cost differences using the posterior distribution at the current time of t=7. Figure 3 presents a similiar picture for bootstrap replications of estimated average policy cost differences using the maximum likelihood estimates at t=7. Comparison of the two figures indicates that the Bayes estimates of average cost difference are much larger and can be positive some of the time; recall that a positive difference implies that it is better never to change. Hence, the Bayes estimates are providing much less evidence of the need to change to the new system. Recall that the changepoint does not occur until time 10. Hence, if the true model parameters were known, the best policy at time 7 would never change.

Figure 11 presents results for current time t=12. Comparing this figure with the corresponding maximum likelihood figure, Figure 5, indicates that there is little practical difference between the two procedures in this case also. However, the widths of the boxes for the Bayesian procedure are larger than those for the

maximum likelihood procedure. The greater widths are an indication of greater uncertainty concerning the future average costs for each policy. As a result, the Bayesian procedure is providing less evidence of the need to upgrade. This greater variability is once again probably due to the Bayes procedure assessing greater uncertainty to the time of onset of degradation. This suggests that the Bayesian procedure may be more cautious than the maximum likelihood procedure.

GLOSSARY

μ: mean number of failures per time period for current subsystem before degradation

η: multiple of degrading trend for current subsystem

σ^2: variance of number of failures per time period for current subsystem

C: time of onset of degrading trend for current subsystem

X_i: number of failures occurring in time period i for current subsystem

λ: mean number of failures per time period for upgraded subsystem

c_0: cost per failure for current subsystem

c_F: initial fixed cost for upgrading current subsystem

c_N: cost per failure for upgraded subsytem

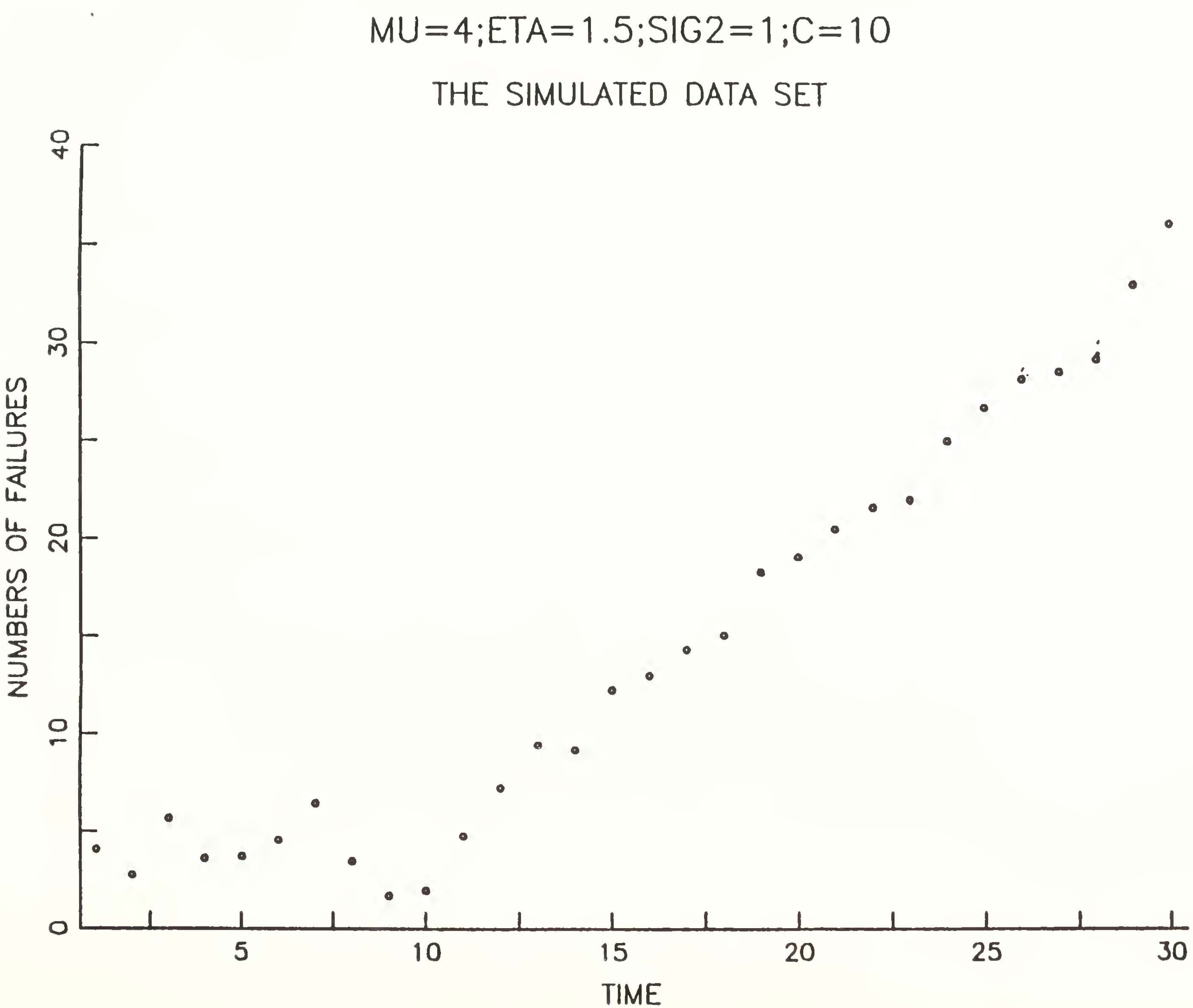

MU=4;ETA=1.5;SIG2=1;C=10
THE SIMULATED DATA SET
NUMBERS OF FAILURES
0
10
20
30
40
TIME
5
10
15
20
25
30

Figure 1

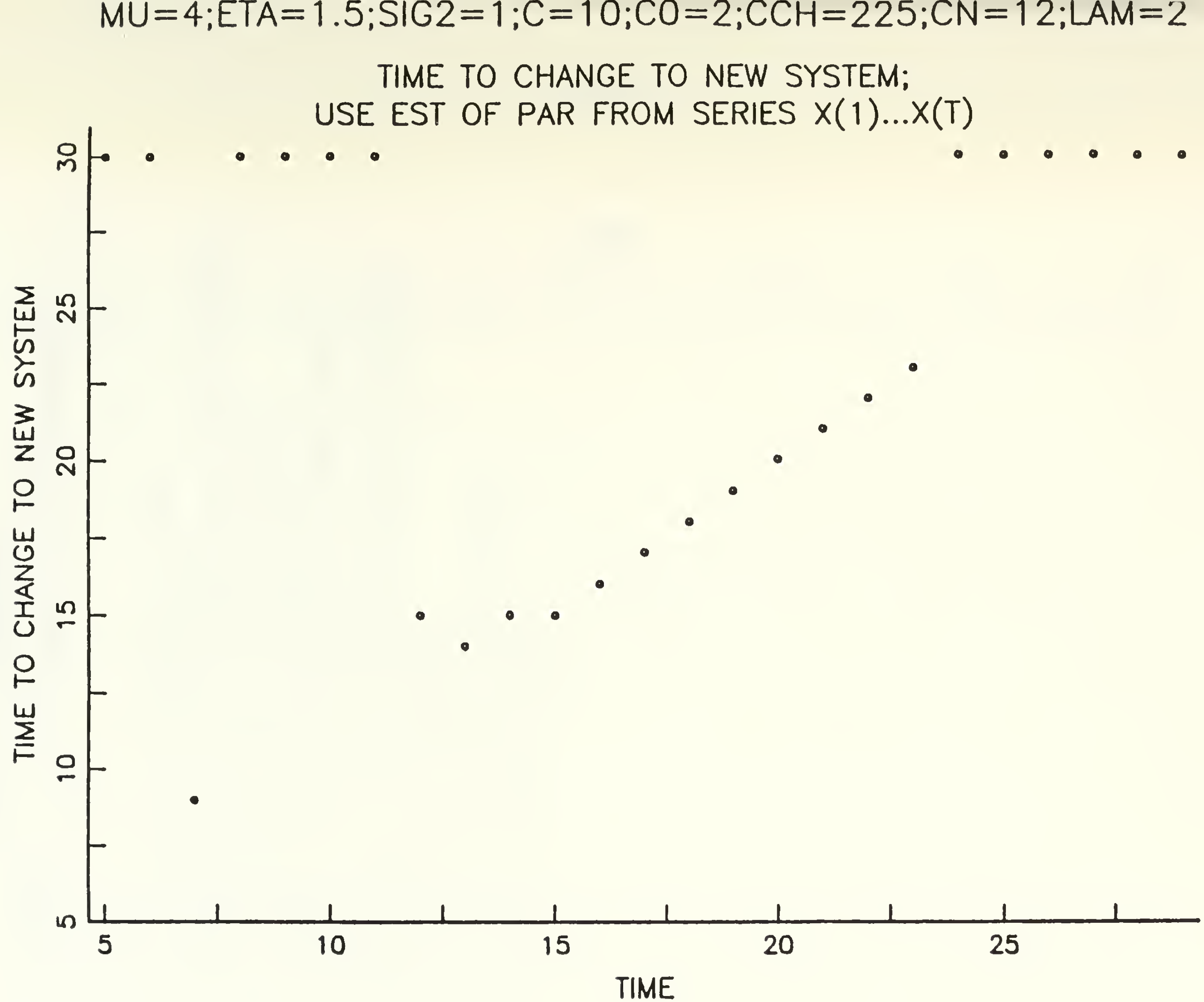
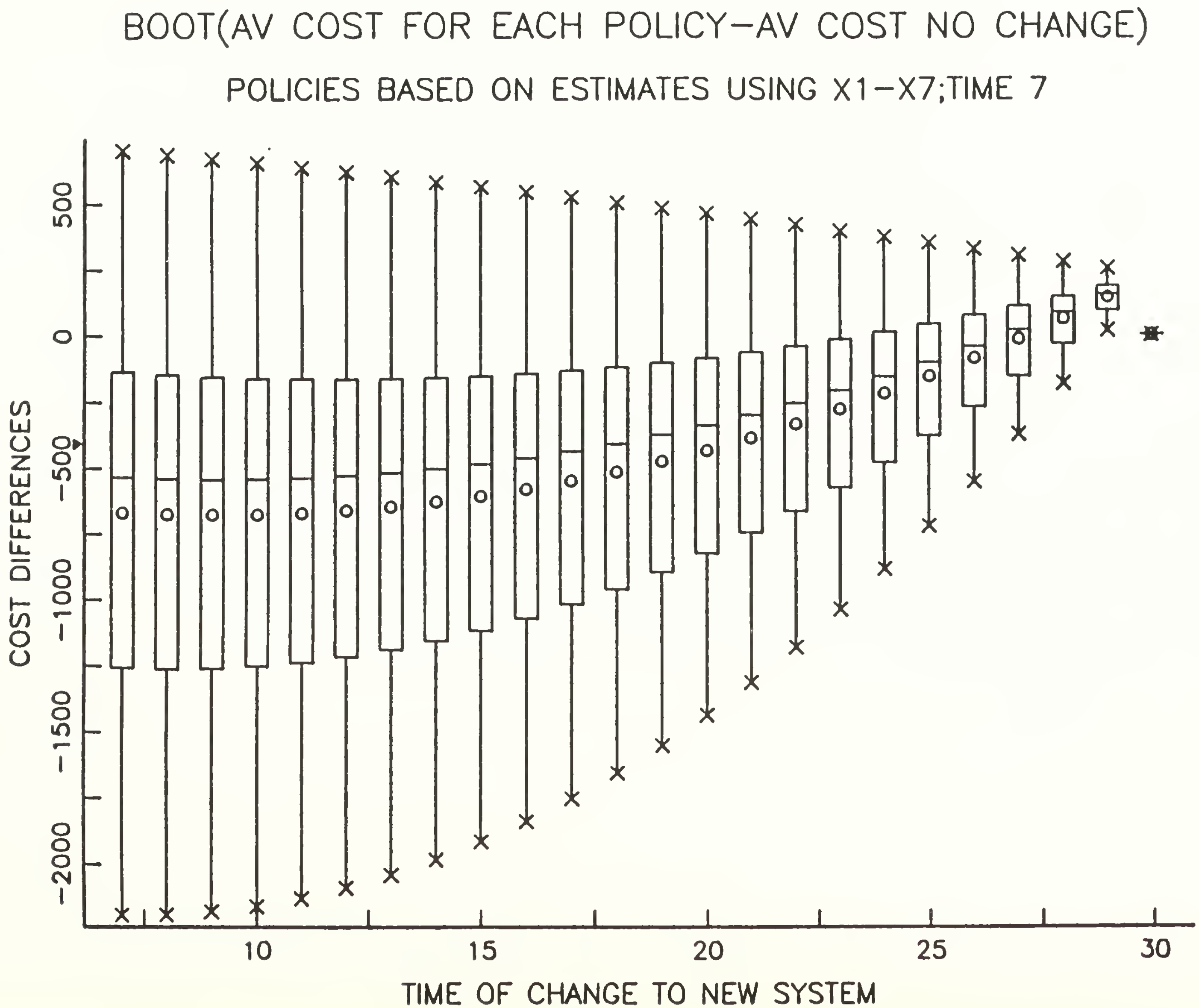

MU=4;ETA=1.5;SIG2=1;C=10;C0=2;CCH=225;CN=12;LAM=2
TIME TO CHANGE TO NEW SYSTEM;
USE EST OF PAR FROM SERIES X(1)...X(T)
TIME TO CHANGE TO NEW SYSTEM
TIME
5
10
15
20
25
30

Figure 2



BOOT(AV COST FOR EACH POLICY-AV COST NO CHANGE)
POLICIES BASED ON ESTIMATES USING X1-X7;TIME 7
COST DIFFERENCES
500
0
-500
-1000
-1500
-2000
10
15
20
25
30
TIME OF CHANGE TO NEW SYSTEM

Figure 3

BOOT(AV COST FOR EACH POLICY–AV COST NO CHANGE)

POLICIES BASED ON ESTIMATES USING X1–X10;TIME 10

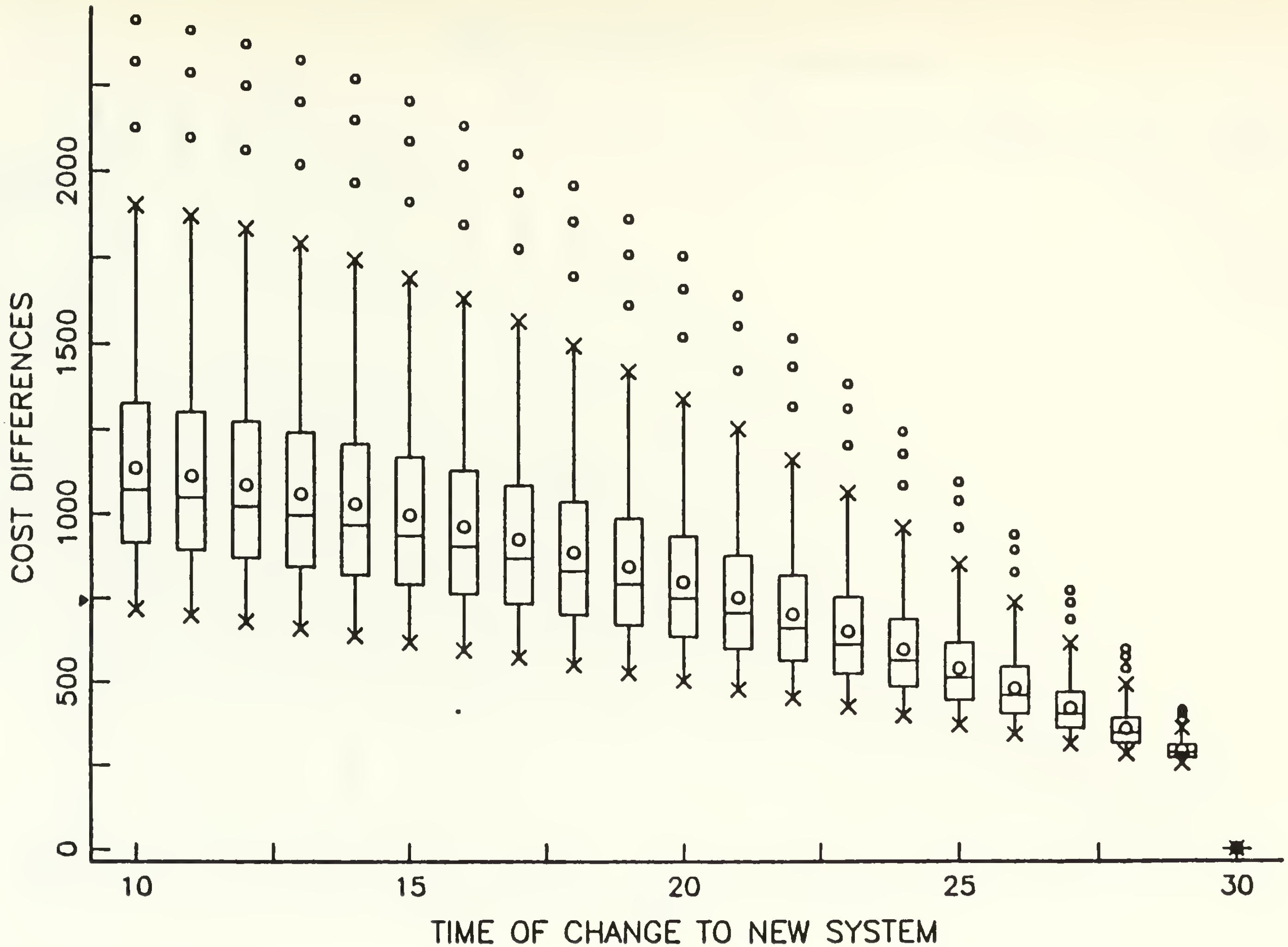


Figure 4

BOOT(AV COST FOR EACH POLICY—AV COST NO CHANGE)

POLICIES BASED ON ESTIMATES USING X1—X12;TIME 12

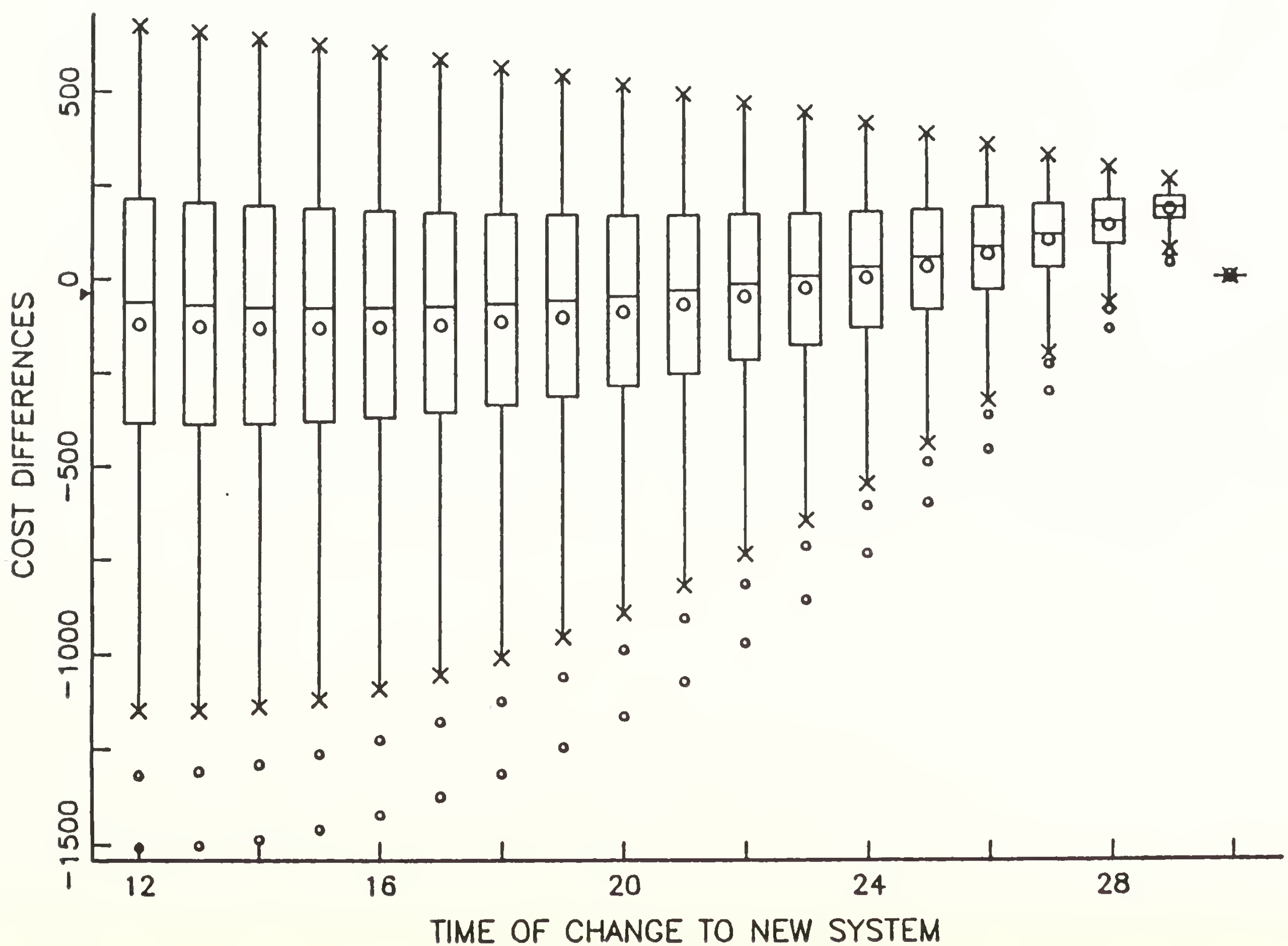
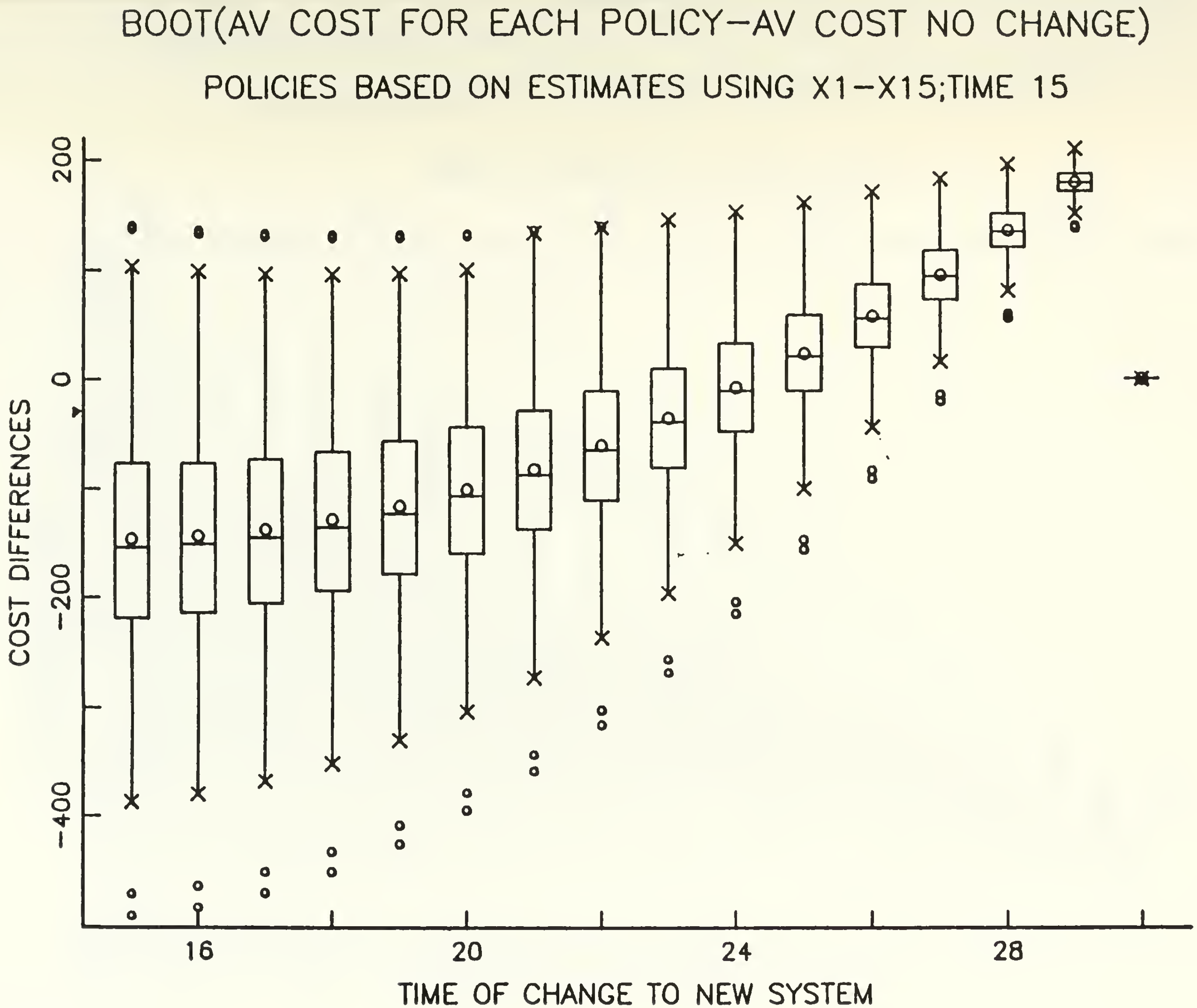


Figure 5



BOOT(AV COST FOR EACH POLICY—AV COST NO CHANGE)
POLICIES BASED ON ESTIMATES USING X1—X15;TIME 15
COST DIFFERENCES
200
0
-200
-400
16
20
24
28
TIME OF CHANGE TO NEW SYSTEM

Figure 6

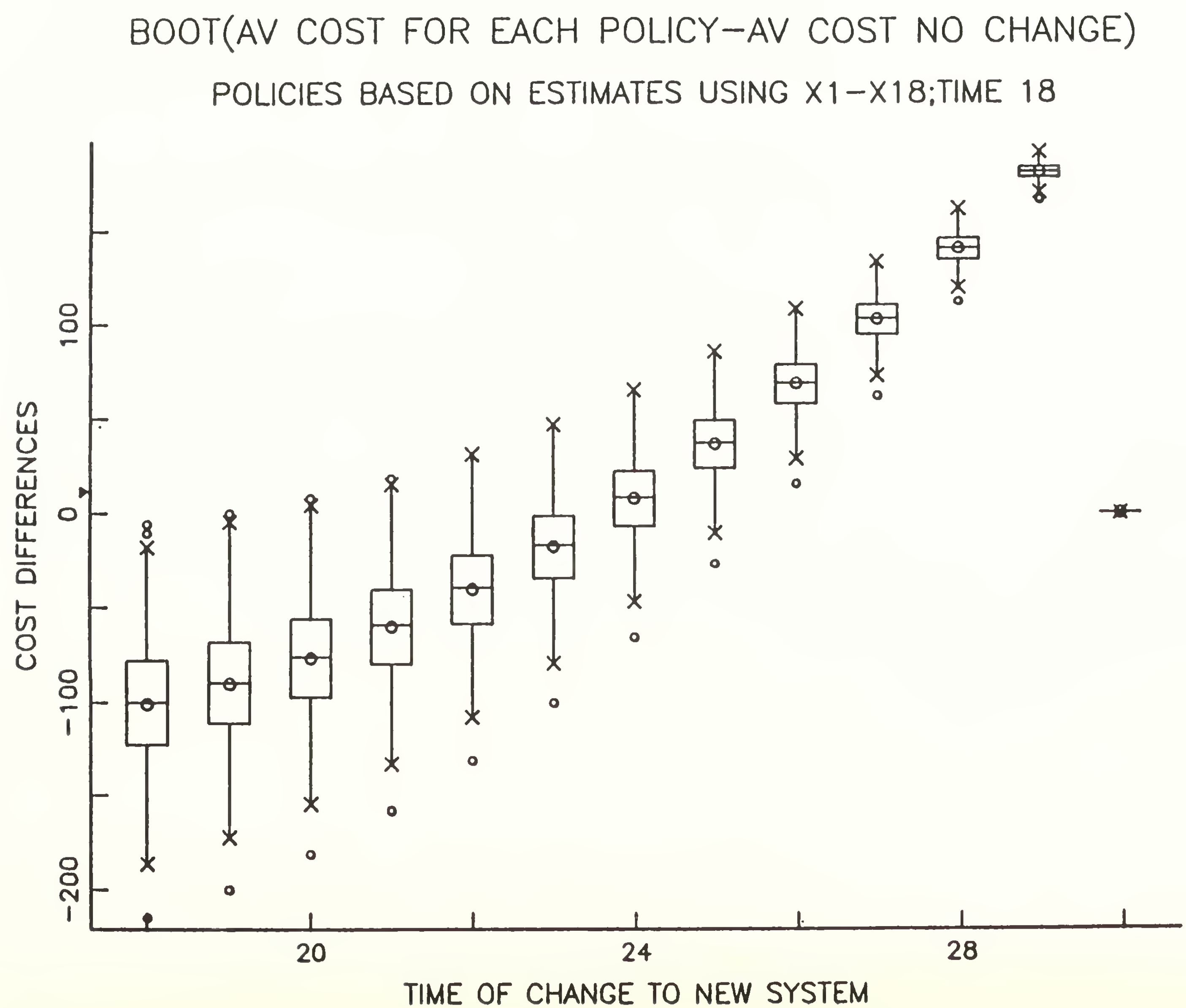

BOOT(AV COST FOR EACH POLICY−AV COST NO CHANGE)
POLICIES BASED ON ESTIMATES USING X1−X18;TIME 18
COST DIFFERENCES
−200
−100
0
100
20
24
28
TIME OF CHANGE TO NEW SYSTEM

Figure 7

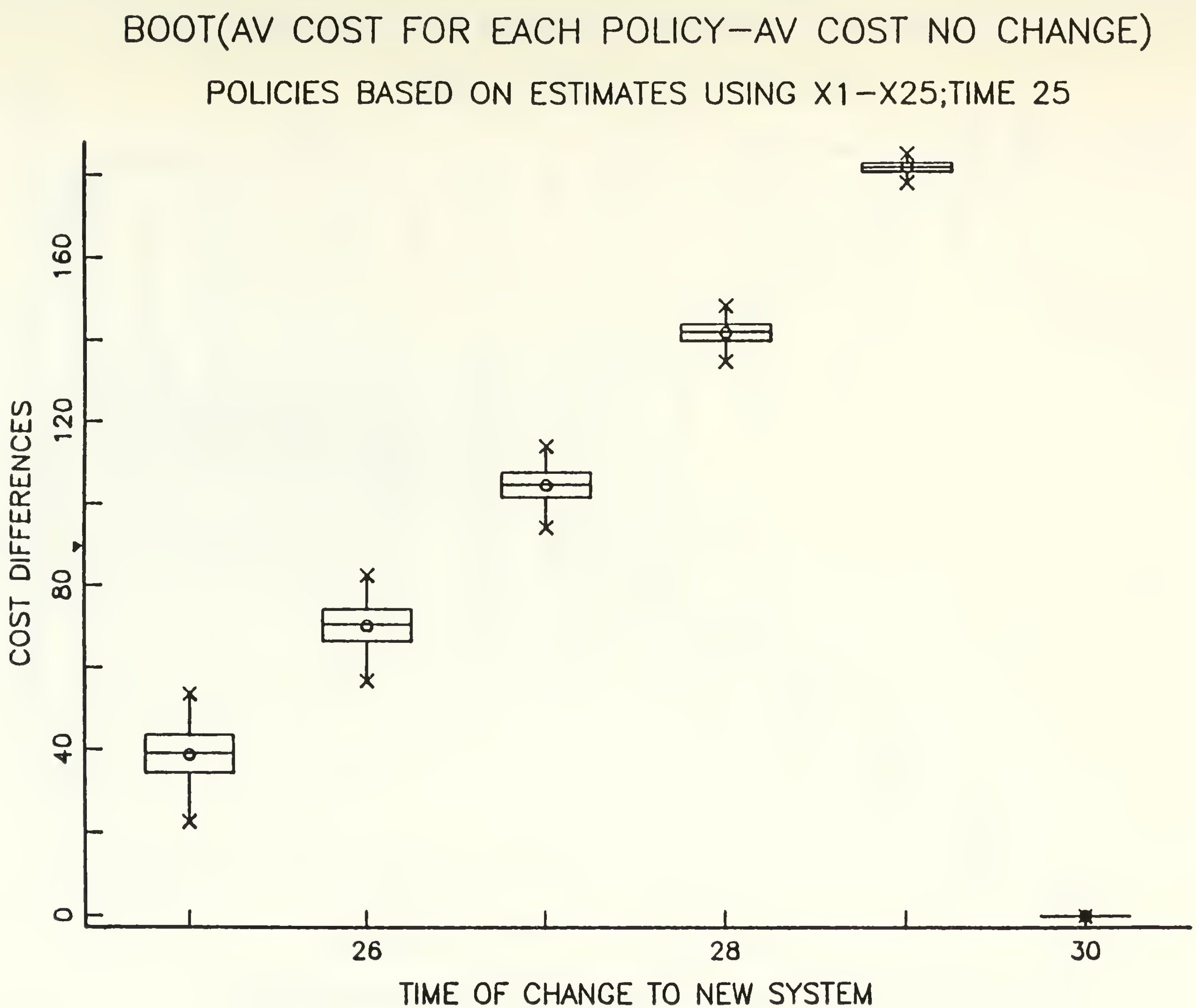

BOOT(AV COST FOR EACH POLICY—AV COST NO CHANGE)
POLICIES BASED ON ESTIMATES USING X1—X25;TIME 25
COST DIFFERENCES
0
40
80
120
160
26
28
30
TIME OF CHANGE TO NEW SYSTEM

Figure 8

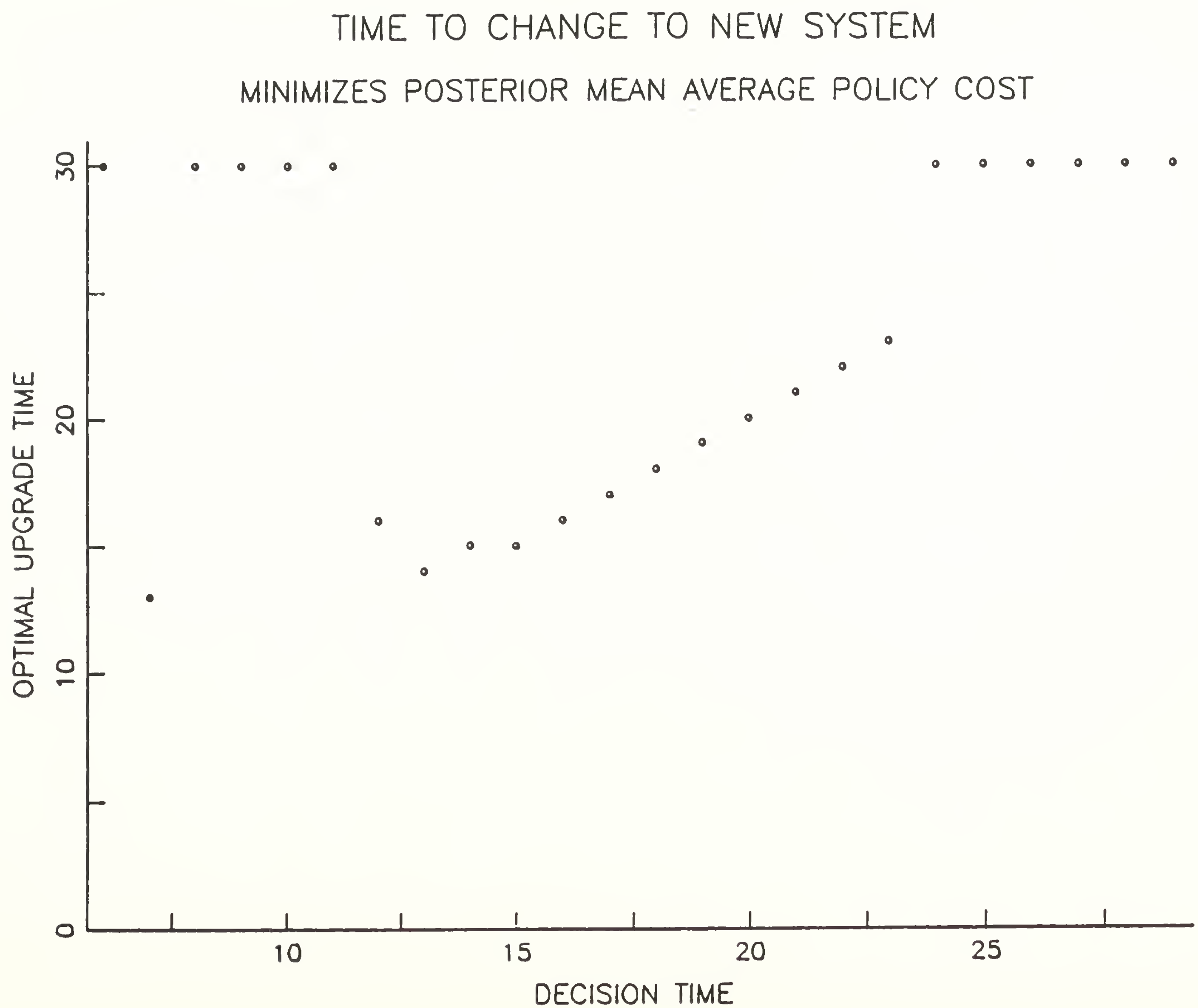

TIME TO CHANGE TO NEW SYSTEM
MINIMIZES POSTERIOR MEAN AVERAGE POLICY COST
OPTIMAL UPGRADE TIME
0
10
20
30
10
15
20
25
DECISION TIME

Figure 9

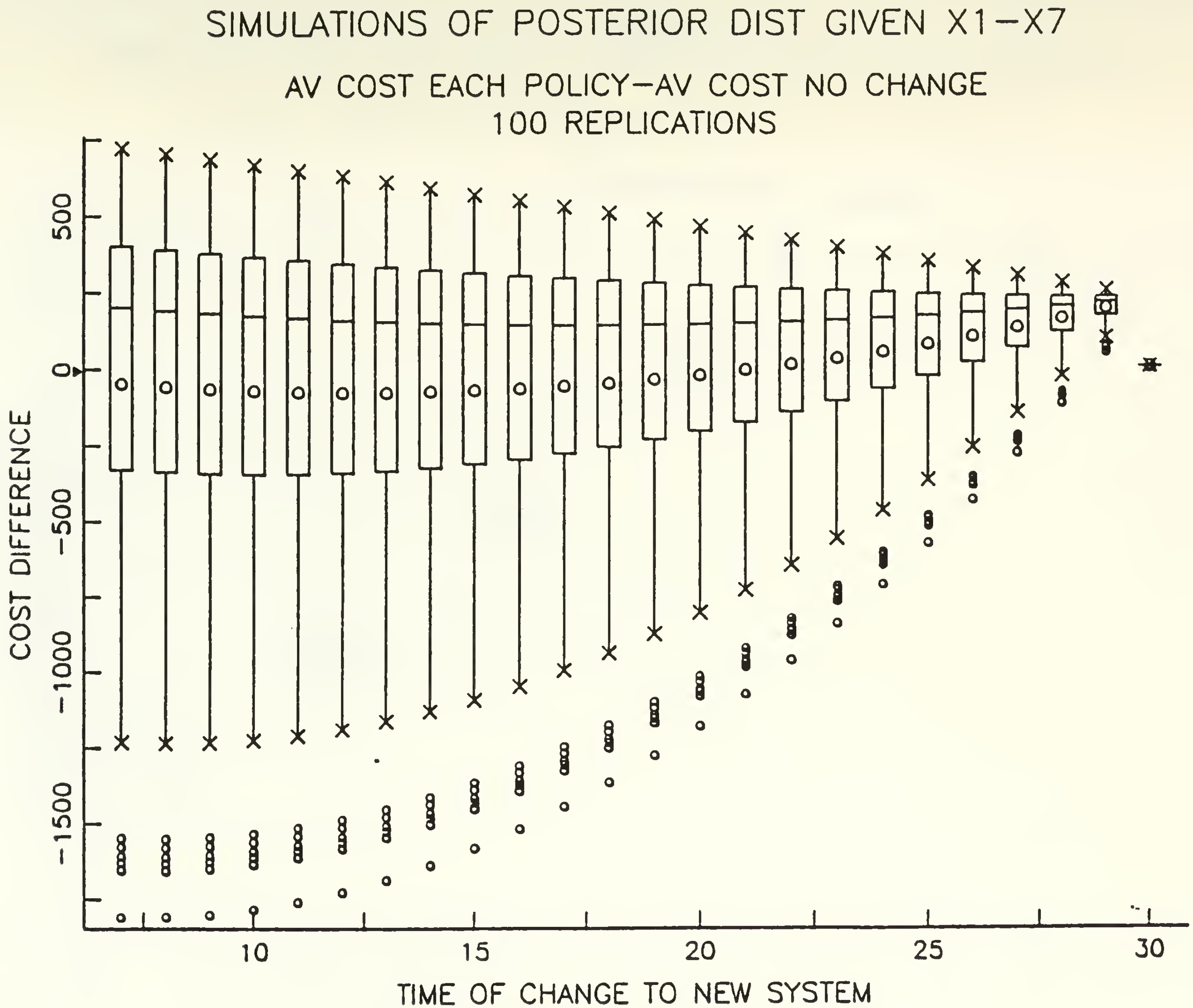
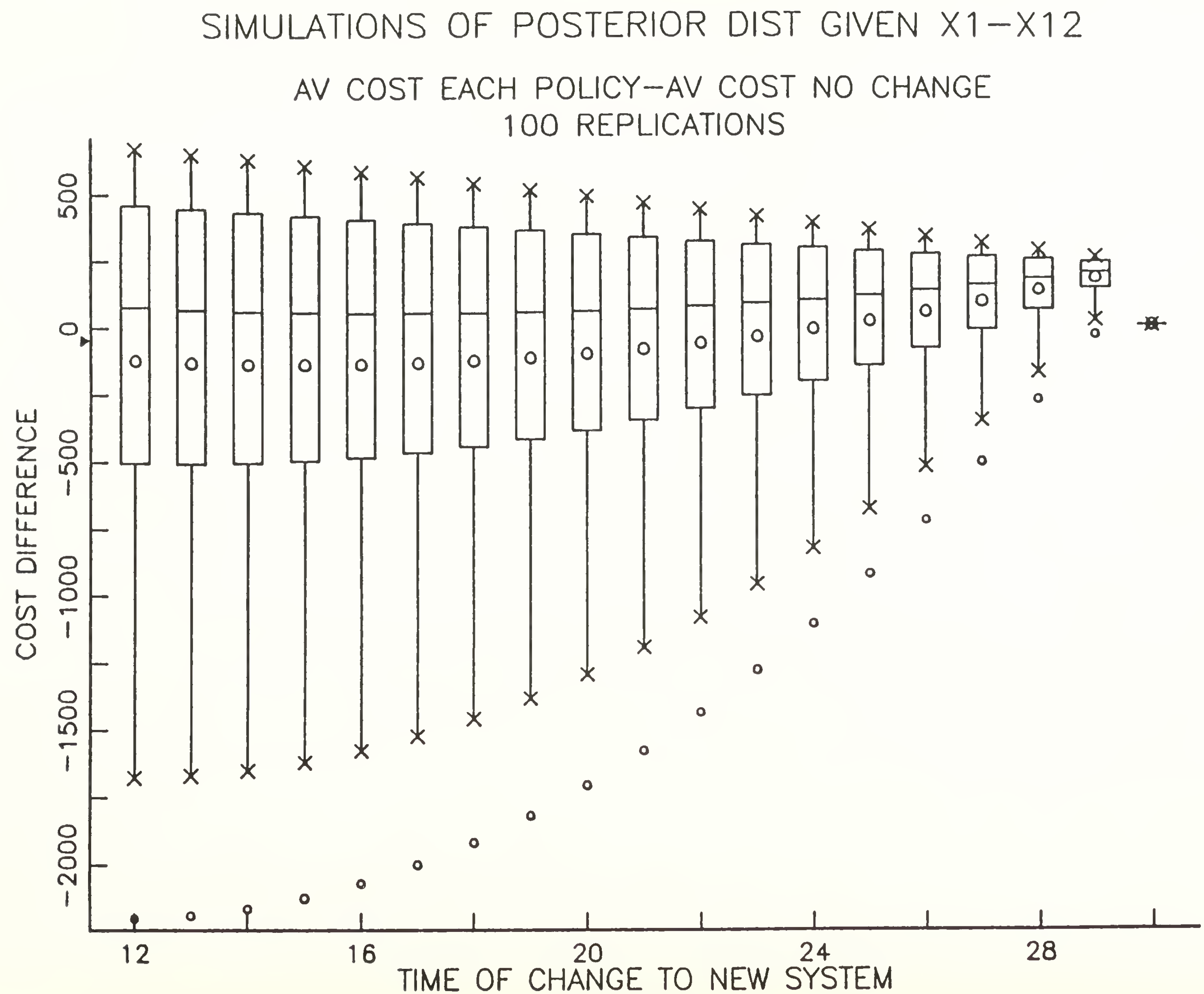

SIMULATIONS OF POSTERIOR DIST GIVEN X1-X7
AV COST EACH POLICY-AV COST NO CHANGE
100 REPLICATIONS
COST DIFFERENCE
500
0
-500
-1000
-1500
10
15
20
25
30
TIME OF CHANGE TO NEW SYSTEM

Figure 10


SIMULATIONS OF POSTERIOR DIST GIVEN X1–X12
AV COST EACH POLICY–AV COST NO CHANGE
100 REPLICATIONS
COST DIFFERENCE
500
0
–500
–1000
–1500
–2000
12
16
20
24
28
TIME OF CHANGE TO NEW SYSTEM
FIGURE 11

INITIAL DISTRIBUTION LIST

1. Library (Code 52)....................2
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Defense Technical Information Center....................2
   Cameron Station
   Alexandria, VA 22314

3. Office of Research Administration (Code 81)....................1
   Naval Postgraduate School
   Monterey, CA 93943-5000

4. Prof. Peter Purdue....................1
   Code OR-Pd
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. Department of Operations Research (Code 55)....................1
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. Prof. Donald Gaver, Code OR-Gv....................5
   Naval Postgraduate School
   Monterey, CA 93943-5000

7. Prof. Patricia Jacobs....................5
   Code OR/Jc
   Naval Postgraduate School
   Monterey, CA 93943-5000

8. Center for Naval Analyses....................1
   4401 Ford Avenue
   Alexandria, VA 22302-0268

9. Dr. David Brillinger....................1
   Statistics Department
   University of California
   Berkeley, CA 94720

10. Prof. Bernard Harris 1
Dept. of Statistics
University of Wisconsin
610 Walnut Street
Madison, WI 53706

11. Prof. W. R. Schucany... 1
Dept. of Statistics
Southern Methodist University
Dallas, TX 75222

12. Prof. D. C. Siegmund 1
Dept. of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

13. Prof. H. Solomon 1
Department of Statistics
Sequoia Hall
Stanford University
Stanford, CA 94305

14. Dr. Ed Wegman 1
George Mason University
Fairfax, VA 22030

15. Dr. P. Welch 1
IBM Research Laboratory
Yorktown Heights, NY 10598

16. Dr. Neil Gerr 1
Office of Naval Research
Arlington, VA 22217

17. Prof. Roy Welsch.. 1
Sloan School
M.I.T.
Cambridge, MA 02139

18. Dr. J. Abrahams, Code 1111, Room 607 1
Mathematical Sciences Division, Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

19. Prof. J. R. Thompson 1
Dept. of Mathematical Science
Rice University
Houston, TX 77001

20. Dr. P. Heidelberger 1
IBM Research Laboratory
Yorktown Heights
New York, NY 10598

21. Prof. M. Leadbetter 1
Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

22. Prof. D. L. Iglehart 1
Dept. of Operations Research
Stanford University
Stanford, CA 94350

23. Prof. J. B. Kadane 1
Dept. of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

24. Prof. J. Lehoczky 1
Department of Statistics
Carnegie-Mellon University
Pittsburgh, PA 15213

25. Prof. M. Mazumdar 1
Dept. of Industrial Engineering
University of Pittsburgh
Pittsburgh, PA 15235

26. Prof. H. Chernoff 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

27. Dr. T. J. Ott 1
Bellcore,
445 South Street
Morristown, NJ 07962-1910
(MRE 2P388)

28. Dr. Alan Weiss 1
AT&T Bell Laboratories
Mountain Avenue
Murray Hill, NJ 07974

29. Prof. Joseph R. Gani..1
Mathematics Department
University of California
Santa Barbara, CA 93106

30. Prof. Frank Samaniego 1
Statistics Department
University of California
Davis, CA 95616

31. Dr. James McKenna...1
Bell Communications Research
445 South Street
Morristown, NJ 07960-1910

32. Dr. Shlomo Halfin..1
Bellcore,
445 South Street
Morristown, NJ 07962-1910
(MRE 2L309)

33. Prof. W. Stuetzle..1
Department of Statistics
University of Washington
Seattle, WA 98195

34. Prof. F. W. Mosteller 1
Department of Statistics
Harvard University
1 Oxford St.
Cambridge, MA 02138

35. Dr. D. C. Hoaglin 1
Department of Statistics
Harvard University
1 Oxford Street
Cambridge, MA 02138

36. Prof. N. D. Singpurwalla 1
George Washington University
Washington, DC 20052

37. Prof. George S. Fishman..1
Curr. in OR & Systems Analysis
University of North Carolina
Chapel Hill, NC 20742

38. Prof. Bradley Efron...1
Statistics Dept.
Sequoia Hall
Stanford University
Stanford, CA 94305

39. Prof. Carl N. Morris..1
Statistics Department
Harvard University
1 Oxford St.
Cambridge, MA 02138

40. Dr. John E. Rolph...1
RAND Corporation
1700 Main St.
Santa Monica, CA 90406

41. Prof. Linda V. Green...1
Graduate School of Business
Columbia University
New York, NY 10027

42. Dr. David Burman...1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

43. Dr. Edward G. Coffman, Jr. 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

44. Prof. William Jewell 1
Operations Research Department
University of California, Berkeley
Berkeley, CA 94720

45. Operations Research Center, Rm E40-164 1
Massachusetts Institute of Technology
Attn: R. C. Larson and J. F. Shapiro
Cambridge, MA 02139

46. Arthur P. Hurter, Jr. 1
Professor and Chairman
Dept. of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60201-9990

47. Institute for Defense Analysis 1
1800 North Beauregard
Alexandria, VA 22311

48. Prof. J. W. Tukey 1
Statistics Dept., Fine Hall
Princeton University
Princeton, NJ 08540

49. Dr. Colin Mallows 1
AT&T Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

50. Dr. D. Pregibon 1
AT&T Bell Telephone Laboratories
Mountain Avenue
Murray Hill, NJ 07974

51. Dr. Jon Kettenring 1
Bellcore
445 South Street
Morris Township, NJ 07962-1910

52. Prof. David L. Wallace 1
Statistics Dept., University of Chicago
5734 S. University Ave.
Chicago, IL 60637

53. Dr. S. R. Dalal 1
Bellcore
445 South Street
Morristown, NJ 07962-1910

54. Dr. B. Doshi 1
AT&T Bell Laboratories
HO 3M-335
Holmdel, NJ 07733

55. Dr. D. M. Lucantoni 1
AT&T Bell Laboratories
Holdmel, NJ 07733

56. Dr. V. Ramaswami 1
MRE 2Q-358
Bell Communications Research, Inc.
445 South Street
Morristown, NJ 07960

57. Prof. G. Shantikumar 1
The Management Science Group
School of Business Administration
University of California
Berkeley, CA 94720

58. Dr. D. F. Daley 1
Statistics Dept. (I.A.S.)
Australian National University
Canberra, A.C.T. 2606
AUSTRALIA

59. Dr. Guy Fayolle 1
I.N.R.I.A.
Dom de Voluceau-Rocquencourt
78150 Le Chesnay Cedex
FRANCE

60. Professor H. G. Daellenbach 1
Department of Operations Research
University of Canterbury
Christchurch, NEW ZEALAND

61. Koh Peng Kong 1
OA Branch, DSO
Ministry of Defense
Blk 29 Middlesex Road
SINGAPORE 1024

62. Professor Sir David Cox 1
Nuffield College
Oxford, OXI INF
ENGLAND

63. Dr. A. J. Lawrence 1
Dept. of Mathematics,
University of Birmingham
P. O. Box 363
Birmingham B15 2TT
ENGLAND

64. Dr. D. Vere-Jones 1
Dept. of Math, Victoria Univ. of Wellington
P. O. Box 196
Wellington
NEW ZEALAND

65. Prof. Guy Latouche 1
University Libre Bruxelles
C. P. 212, Blvd. De Triomphe
B-1050 Bruxelles
BELGIUM

66. Capt J. Gill 1
NAVAIR (AIR 419)
Washington, D.C. 20550